Structural and Computational Studies of Small Organic and Biological Molecules

Patricia Lozano-Casal

A Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy to the School of Chemistry, University of Edinburgh.

February 2006

Declaration

I declare that this thesis was written by myself and that the work detailed in this thesis is my own, or I have contributed substantially to such work, except where specific reference is made to the work of another.

Patricia Lozano-Casal

ABSTRACT

Over the last three decades high-pressure X-ray diffraction techniques have been widely utilised to perform structural studies in many areas of research. For example, physicists make use of these experimental techniques to investigate metals, conductor and semi-conductor compounds among others, whereas geochemists apply them to study the conditions deep within the Earth's interior. Furthermore, pressure studies have reached an important status in chemistry, biology and planetary science, and have proved to be a new notable tool to study the structure of a variety of small molecule compounds, from inorganic (e.g. rock salt) to organic (e.g. urea) and biological molecules (e.g. amino acids). The main reason for this is the necessity to obtain a better understanding of different processes which take place at extreme conditions of pressure, such as the existence of life in the deep ocean (e.g. extremophiles) or the finding of amino acids, such as cysteine, in space. Small molecules, such as glycine and glutathione, may play important roles in these biological processes and therefore a good knowledge of their structural features could be essential to explain how they happen.

The work described here focuses on one of the principal features of intermolecular bonding in small organic and biological molecules, namely the hydrogen bond. Although the hydrogen bond has been studied for almost a century, this important directional intermolecular interaction is still one of the most highly investigated topics in chemistry and biology due to the many aspects of its nature which are still unknown or not well understood. The hydrogen bond is directly involved with the crucial biochemical processes of amino acids, peptides, proteins and even DNA and RNA and therefore, in short, it is essential for life. Its importance goes beyond biomolecules and organic molecules, since the hydrogen bond is also present in inorganic compounds, such as clusters containing different types of ligands and metals, e.g. manganese and iron, and consequently it is implicated in many catalytic pathways.

In this work, we have combined experimental and computational techniques to investigate the effect of pressure on the crystal structures of amino acids and small organic molecules, and in particular on the crystal packing of cyclopropylamine, α-glycine, L-α-aspartic acid and L-α-glutamine. The experimental study of these structures was aimed to determine the principal structural changes as a function of pressure. The computational study was carried out to investigate the energetics of the crystal structures: i.e., the sublimation, lattice and proton transfer energies, as well as the energies of the individual hydrogen bonds, in order to relate the structural changes observed by experiment. A new computational method was initially tested on the simplest naturally occurring amino acid, α-glycine, which had been previously studied experimentally at high pressure by our research group.

Experimental results support the existence of a new polymorph of cyclopropylamine (Phase II), which was found at 1.2 GPa. This new molecular structure crystallises in the orthorhombic space group *Pbca*, with 8 molecules in the unit cell and one in the asymmetric unit. The molecular packing is formed by zig-zag chains of molecules linked via a short N-H…N hydrogen bond, with unusually only one of the hydrogen atoms of the NH_2 group of the cyclopropylamine molecule involved in the hydrogen bonding network. A computational study was applied to the two polymorphs of cyclopropylamine (Phase I and Phase II), in order to investigate the energetics of these systems. It was found that the energies of the hydrogen bonds present in the Phase I and Phase II crystal structures of cyclopropylamine have very similar energies, falling in the range of 1 to 4 kcal mol^{-1} (6 to 16 kJ mol^{-1}).

Experimental results on amino acids showed that the effect of pressure (from 0 to approximately 6 GPa) did not cause significant changes in the crystal structures of L-α-aspartic acid and L-α-glutamine, other than a decrease in the lengths of the various hydrogen bonds and intermolecular contacts. Additionally, *ab initio* calculations carried out on α-glycine, L-α-aspartic acid and L-α-glutamine indicated that the hydrogen bonds present in the crystal structure of amino acids exhibit very different energies, falling over a much wider range of 1 to 30 kcal mol^{-1} (~4 to

120 kJ mol^{-1}), compared to that found for small organic molecules. This finding does not agree with what was expected according to previous hydrogen bonding classifications, and we attribute this to the zwitterionic nature of the amino acid molecules which is present in the solid state. Finally, it was found that the energies of the hydrogen bonds found in the amino acids investigated correlates with the experimental compressibility studies, so that the hydrogen bond exhibiting the lowest energy is also the weakest, and most easily compressed.

Acknowledgements

I would like to thank my supervisors Dr David Allan and Dr Carole Morrison. First of all, for giving me the chance to do this exciting PhD. Secondly, for all their help and advice over the last three years. Not only have they taught me about Crystallography and Computational Chemistry, but helped me to become more independent and to appreciate the mysteries of Chemistry even more. Finally, I would like to thank them both for all the support that they have offered me, giving me strength to go on during testing times. For all of this, I shall always be grateful to you.

I would also like to thank:

Dr. Simon Parsons, undoubtedly the best crystallographer I have ever met, for teaching me about one of the hardest topics that I have ever come across, which is crystallography. He has made everything so easy to understand and remember … How could I ever forget about Symmetry??

Dr. Colin Pulham, for sharing his knowledge of Chemistry and for showing me the type of scientist that one day I wish to become.

Dr. Alice Dawson and Dr. Iain Oswald, for all their help and friendship!

Francesca, Stephen, James, Fraser and Pete, for being the best group-mates than anyone can dream of. Thank you guys for being such good friends and for all the jokes and chats that made the office a great place to work in!! I can still remember Fran's screams when she thought someone had stolen her conference poster!!!

Thanks also to Alistair, Chris, David Manuel, Laura, Murshed and Martin for allowing me to get to know them and laugh with them at conferences and pubs, they are great fun! Who said that Chemists and Physicists were boring??

Rob for being there for me since I came to study to Edinburgh for the first time. Five years have nearly gone by and you are still my friend!! I still think we should form that rock band …

I would like to thank my Spanish friends Chus (Chusalinda), Sara, Iria, Carmela and Mónica for making my holidays so enjoyable with all those night outings and the time we spent together. Also for being so close to my heart.

Leticia, for being my best friend. Who else would be able to keep up with all my crazy behaviour over the last three years?

Joe, for being one of the cleverest guys I have ever known! You left me exhausted after every discussion … religion, music, literature, etc … thanks for keeping my mind busy and well educated!!!

My "gym buddies", Nadia, Lindsay, Pomme, Sarah, Marcos and JC for being so good to me and for keeping my body and soul fit!!

Aimie for being such a good flatmate and friend to me. I could not have done it without all her support and … I hope one day I will do the same for you!! You are the best!

My sisters Tabby (THM) and Ana (Dino Malo!!) for being the best sisters I could ever wish for and for the long telephone chats about everything apart from Chemistry!!

My brother-in-law, John, for all the scientific (and less scientific) conversations.

My nephew and niece, James and Katie, for their innocence and for letting me play with them for hours and hours.

My dog Fili and my "white bitch" Gusa, for being so cuddly and loveable.

José and Marisa, for being the best parents in the whole world; for making me the person I am today; for supporting me in every decision I have ever made, especially choosing a career in Chemistry abroad (I shall always remember Dad trying to convince me to do Literature or Architecture instead) and, most importantly, for believing in me. I am truly proud of being your daughter and I shall always love you.

Last but not least, to those who have not been mentioned here but who supported me throughout my life -teachers, friends and family - and for that, they will always occupy a special place in my heart.

This thesis is dedicated to my parents, José and Marisa, the best two people in the whole world, whom I deeply admire and love.

Contents

Chapter 1

Introduction

1.1 Introduction

The structure of molecular systems in the solid state is, in most cases, dominated by the formation of intermolecular interactions between the constituent molecules. The nature of these interactions varies from strong quasi-covalent interactions, such as those present in the combination of acids with their conjugate bases, to weak van der Waals interactions, like those present in methane (CH_4). One of the most common and interesting types of interactions is the hydrogen bond, which is particularly important in biological molecules, since it stabilises the different structures of proteins and even DNA.

Pressure and temperature are fundamental thermodynamic variables, which can be used to change intermolecular distances and probe various interactions. Pressure has a profound effect on intermolecular distance and consequently can alter the hydrogen bond significantly as it is one of the softest intermolecular interactions. Thus, when pressure is applied to a molecular material the geometry of the hydrogen bonding is affected, which in turn may lead to structural phase transitions and the formation of new polymorphs.

X-ray diffraction is one of the few structural techniques that can give accurate information about how molecules pack together in the crystalline state giving a unique solution, whereas other techniques, such as solid-state NMR, Raman or IR spectroscopies, require the combination of two or more techniques and rather difficult interpretation to obtain a solution for the crystal structure. Thus, X-ray diffraction techniques can be used to follow changes in the structure of crystalline materials. On the other hand, computational techniques, such as Quantum Mechanics, enable the study of the solid state from an energetic point of view, as they can calculate thermodynamic properties, such as sublimation and lattice energies, which are difficult or impossible to determine experimentally. In addition, the use of computational techniques might be of help to complete models taken from the high-pressure X-ray diffraction experiments (e.g. locating hydrogen atom positions), which may be in pressure/temperature regimes that are difficult to access *via* neutron-diffraction techniques (which can locate H-atom positions directly). Thus, the use of

X-ray diffraction and computational techniques can be of help to study the effect of pressure on crystal structures and it is the combination of both techniques that gives a complete view of the nature and properties of the hydrogen bond and crystal packing.

1.2 The nature of the hydrogen bond

In 1920 Latimer and Rodebush suggested that "a free pair of electrons on one water molecule might be able to exert sufficient force on a hydrogen held by a pair of electrons on another water molecule to bind the two molecules together" (Latimer et al., 1920). This speculation (greeted with some scepticism at the time) initiated two generations of research into the hydrogen bond, the output of which has been prodigious (Taylor *et al.*, 1984a).

1.2.1 Definition

The first modern definition of the hydrogen bond was given by Pimentel *et al.* (1960), who wrote that "A hydrogen bond is said to exist when (1) there is evidence of a bond, and (2) there is evidence that this bond sterically involves a hydrogen atom already bonded to another atom". However, the description of the hydrogen bond made by Pimentel *et al.* does not give any hint of the nature, in terms of electronegativity and polarity, of the atoms forming the interaction, or the geometry restrictions of the hydrogen bond. Thus, two more detailed definitions of the hydrogen bond were given by Steiner *et al.* (1993, 2002). Steiner considered a hydrogen bond as "any cohesive interaction X-H…A where H carries a positive and A a negative (partial or full) charge and the charge on X is more negative than on H" and also proposed that: *An X-H…A interaction is called a "hydrogen bond", if 1. it constitutes a local bond, and 2. X-H acts as a proton donor to A.* The geometry restrictions which must be applied to the X-H…A interaction in order to classify it as a weak or strong hydrogen bond were first given by Emsley in 1981 in his paper on "Very strong hydrogen bonding", and which is presented in Table 1.1.

In terms of denomination, during the production of this thesis, the X-H…A nomenclature will be substituted by D-H…A, where D, H and A are the donor, hydrogen and acceptor atoms respectively. However, in the tables taken from the literature the original terminology will remain unchanged.

Property	*Weak hydrogen bonding* A-H…B	*Strong hydrogen bonding* A-H-B
Bond length r(A…B)	Slightly less than the sum of the van der Waals radii of A and B.	Significantly less (>30 pm) than the sum of the van der Waals radii of A and B
Location of H	Near its parent atom: r(A-H) ≈ r_{cov} (A-H)	Centred, or roughly so, but not covalently near to A or B
Bond vibrational mode ν_{AH}	Broad bands and shifted to lower frequencies; still in 2000-3000 cm^{-1} region	Very broad bands in region below 1600 cm^{-1}
$\Delta\nu_{AH\ (bonded)}/\nu_{AH\ (non\text{-}bonded)}$ $\nu_{AH}/\nu_{AD\ (isotope\ frequency\ ratio)}$	< 25% 1.35 or less	> 25% Tends to 1 but may be > 1.35
Bond energy	E(A-H…B) < 50 kJ mol^{-1} And most < 30 kJ mol^{-1} (measured)	E(A-H-B) > 50 kJ mol^{-1} And some > 100 kJ mol^{-1} (calculated)
Proton shielding; chemical shift	Slight downfield shift from $\delta(^1H,AH)$ non-hydrogen-bonded	Large downfield shift, sometimes below 20 ppm

Table 1.1: Weak and strong hydrogen bonding (Emsley, 1981).

1.2.2 Constituent interactions

The D-H…A hydrogen bond is composed of three atoms (i.e. donor, hydrogen and acceptor), which means that the hydrogen bond is then a group property (Desiraju *et al.*, 1999). Consequently, the hydrogen bond is not a simple interaction but a complex one, formed by at least five constituents of different natures (Steiner, 2002).

Following the Morokuma (1971, 1977) theoretical partitioning of the total hydrogen bond interaction, the overall hydrogen bond energy (E_{tot}) can be split into contributions from electrostatics (E_{elec}), charge transfer (E_{ct}), polarisation (E_{pol}) dispersion (E_{disp}), and exchange repulsion (E_{er}). Of these terms, only the exchange repulsion is repulsive, whereas the others are attractive at all distances. The different contributions can be classified in directional and non-directional terms. The non-directional terms are exchange repulsion and dispersion, which can be grouped to form the "van der Waals interaction". The strength of these contributions depends on their distance and angular characteristics. The electrostatic term, for example, is directional, of long range and diminishes as $-r^{-3}$ for dipole-dipole, $-r^{-2}$ for dipole-monopole, and $-r^{-1}$ for monopole-monopole interactions. Polarisation decreases faster than the electrostatic term, as $-r^{-4}$, whereas the charge-transfer term decreases even faster, following e^{-r}. The dispersion term is isotropic with a distance dependence of $-r^{-6}$, whereas the exchange repulsion increases exponentially. All these interactions present different weights depending on the hydrogen bond and the nature of the atoms forming the interaction. Electrostatics is dominant in strong hydrogen bonds, where it can contribute as much as 60-80% of the attractive terms (Desiraju *et al.*, 1999); in "normal" hydrogen bonds, however, it still remains the most important term, together with charge-transfer, which is also present. In weak hydrogen bonds, the van der Waals terms become very important and might contribute as much as electrostatics to the total hydrogen bond energy (Steiner, 2002).

1.3 The geometry of the hydrogen bond

1.3.1 Geometry and classification

The geometry of the D-H…A-Y hydrogen bond, where the interaction is extended on the acceptor side, can be defined in terms of five parameters: the distances (1) D-H (r), (2) H…A (d), (3) D…A (R), and the angles (4) D-H…A (θ), and (5) D-H…A-Y (ϕ) (Figure 1.1). Many studies have been carried out to investigate the factors influencing hydrogen bond distances, angles and lone-pair directionality.

Some of these were done by Brown (1976) and Taylor *et al.* (1984a, 1984b, 1984c). The classification of the hydrogen bond in terms of its strength can be made regarding the geometry of the D-H…A interaction. Thus, the classification in Table 1.1 can be improved taking into account the geometry restrictions of the hydrogen bond, giving rise to Jeffrey's classification (1997), which can be found in Table 1.2 (Steiner, 2002).

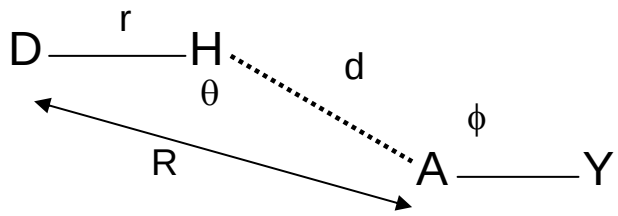


Figure 1.1: Geometrical parameters for a hydrogen bond (*r*, *d* and *R* are distances D-H, H…A and D…A respectively, whereas θ and ϕ are the angles D-H…A and D-H...A-Y).

Table 1.2 shows the classification of the hydrogen bond in terms of strong (covalent), moderate (electrostatic) and weak (electrostatic/dispersion) interaction, depending on the geometry of the interaction (Gilli *et al.*, 2000). However, this classification was based on compounds (dimers) studied in the gas phase, so some differences will appear when using this information to investigate solid-state hydrogen bonds, where molecules are rigidly held in place vibrating about a fixed point and linked by intermolecular forces (i.e. hydrogen bonding), whereas in the gas phase they move freely and lack any additional intermolecular contacts as molecules are in an isolated environment.

The existence of very strong hydrogen bonds, which are quasi-covalent, was described by Emsley (1981), who studied homonuclear (e.g. O-H…O) and heteronuclear (e.g. F-H…O) hydrogen bonds in many different systems, such as difluorides, dicarboxylates and fluoride-water. Experimentally, X-ray and neutron diffraction can be used to study the symmetry of the hydrogen bond, which has been proposed to increase the strength of the interaction, compared to less symmetric or asymmetric hydrogen bonds (Emsley, 1981). This explains why homonuclear

hydrogen bonds are stronger than heteronuclear hydrogen bonds (Gilli *et al.*, 2000). Other strong hydrogen bonds studied were O-H…N, which has great significance in biological systems, and F-H…N among others. Perrin *et al.* (1997) define strong hydrogen bonds as a special class of hydrogen bond characterised by "great strength, short distances, a low or vanishing barrier to hydrogen transfer, and distinctive features in the NMR spectrum". Consequently, this description fully agrees with Emsley (1981) and Jeffrey's (1997) classifications previously mentioned. Strong hydrogen bonds are especially important in nature, as they are the key to obtaining a better understanding of the structure and properties of many biological compounds, such as water, proteins or even DNA. In addition, strong hydrogen bonds have been found to be important in many enzyme-catalysed reactions.

Property	*Strong*	*Moderate*	Weak
Interaction type	Strongly covalent	Mostly electrostatic	Electrostatic/ Dispersion
Bond lengths [Å] H…A	1.2-1.5	1.5-2.2	>2.2
Lengthening of X-H [Å]	0.08-0.25	0.02-0.08	<0.02
X-H *versus* A…H	X-H ≈ H…A	X-H < H…A	X-H << H…A
X…A [Å]	2.2-2.5	2.5-3.2	>3.2
Directionality	Strong	Moderate	Weak
Bond angles [°]	170-180	>130	>90
Bond energy [kcal mol^{-1}]	15-40	4-15	<4
Relat. IR shift $\Delta\nu_{XH}$ [cm^{-1}]	25%	10-25%	<10%
^{1}H Downfield Shift	14-22	<14	

Table 1.2: Strong, moderate, and weak hydrogen bonds following the classification of Jeffrey (1997). The numerical data are guiding values only (taken from Steiner, 2002).

Weak hydrogen bonds, such as C-H…O, are less common than strong or moderate hydrogen bonds and their study is only comparatively recent. Nevertheless, the weak hydrogen bond is considered as an interesting and important type of interaction due to its presence in biological systems, such as proteins and DNA. There are many experimental and computational studies on the weak hydrogen bond, which attempt to give a better description of its role in nature. Some of these studies were carried out by Desiraju *et al.* (1999) and Scheiner *et al.* (2001).

1.3.2 Energies

The energy of hydrogen bonds in the solid state cannot be obtained directly, and in isolation by experiment. However, if a good structural model of the molecular system is available through the experiment, computational techniques can be used to calculate the energy of the hydrogen bonds present in the structure. Thus, many theoretical studies have been carried out on a number of different systems, especially in the gas phase (Del Bene *et al.*, 2001; Tsuzuki *et al.*, 2001). Only recently, the study of the hydrogen bond energy has been extended to the solid state (Fortes *et al.*, 2003; Morrison *et al.*, 2003, 2004, 2005).

From Table 1.2 it can be seen that, in general, the hydrogen bond energy falls in the range of 0.2 and 40 kcal mol^{-1} (i.e. 0.8 and 168 kJ mol^{-1}). The different types of hydrogen bonds are expected to have certain values of hydrogen bond energies. Thus, strong hydrogen bond energies will fall in the range 15-40 kcal mol^{-1} (i.e. 63-168 kJ mol^{-1}), whereas moderate hydrogen bonds will present energies between 4-15 kcal mol^{-1} (i.e. 17-63 kJ mol^{-1}). Finally, weak hydrogen bonds will have energies lower than 4 kcal mol^{-1} (i.e. < 17 kJ mol^{-1}). In the solid state the hydrogen bond is taken out of its optimal geometry due to the need for an effective packing of the molecules involved in the crystal. Apart from the hydrogen bond there are many other effects arising from the close proximity of the molecules, which will affect the energies of the interactions. In terms of the hydrogen bonds present in zwitterionic amino acid systems, unusual energy values will be calculated and presented in this thesis. Energy values will vary from 1 to 30 kcal mol^{-1} (4 to 125 kJ mol^{-1}).

1.3.3 Donor and acceptor directionalities

An important characteristic of hydrogen bonds is their directionality (i.e. the angle θ, formed by the D-H…A atoms), as it is this preference for linearity that distinguishes the hydrogen bond from van der Waals interactions. Linearity is preferred because it optimises the electrostatic interaction between the constituent atoms of the hydrogen bond. There is a certain correlation between the hydrogen bond distances and the linearity of the interaction. Thus, the closer the angle θ is to 180°, the shorter the distance becomes. The angle θ falls in the range 170-180° for strong hydrogen bonds, between 130 and 170° for moderate hydrogen bonds, whereas if the angle θ takes values between 90 and 130° the hydrogen bond will have a weak character (Table 1.2). The degree of directionality depends on the polarity of the donor atom; the more polar the donor is, the more linear the interaction becomes (Steiner, 2002). Hydrogen bonds are directional also at the acceptor side. The linearity of the different hydrogen bond types can be studied by doing a search on the Cambridge Structural Database (CSD) and plotting distributions of the results, as carried out by Steiner *et al.* (1996) and Hay *et al.* (2002). Steiner *et al.* reported that the acceptors of hydrogen bonds generally exhibit a more complex directionality than donors, which depend on their chemical nature. They showed that the acceptor directionality in unhindered C-H…O=C hydrogen bonds donated by acidic C-H groups is very soft, with a significant preference for the conventional carbonyl lone-pair direction. When investigating the oxygen acceptor directionality in oxyanion hydrogen bonds, Hay *et al.* established for the first time the existence of a significant and general oxygen acceptor directionality in hydrogen-bonded complexes with trigonal planar and tetrahedral oxyanions. Finally, the dependence of the bond directionality has not only being investigated through CSD surveys but also theoretically. Ireta *et al.* (2004) investigated the degree of agreement of calculated geometry parameters with the experimental data available in the literature, testing a number of different computational methods. It was shown that the agreement is better for linear hydrogen bonds, whereas for those deviating from a linear arrangement the accuracy of some methods decrease. Thus, it is possible to use the donor and acceptor directionality in the hydrogen bond to test computational methods.

1.3.4 Multifurcated hydrogen bonds: classification

One of the terminologies concerning hydrogen bond formation most used in the literature includes the terms two-centred, three-centred and four-centred hydrogen bonds (1, 2 and 3 in Figure 1.2) (Jeffrey *et al.*, 1982). This nomenclature considers that the terms "linear", "bifurcated" and "trifurcated" are not sufficient to describe hydrogen bonds, since two categories of hydrogen bonds could be named as "bifurcated" (2 and 4 in Figure 1.2).

Hydrogen bonds with more than three acceptors are possible, although they are very rarely found in crystal structures, due to the acceptors´ high special density requirements. The geometry of N-H…O=C three-centred and four-centred hydrogen bonds was investigated by Taylor *et al.* (1984c), whereas Jeffrey *et al.* (1984) studied three-centre hydrogen bonds in the crystal structures of amino acids, formed by the NH_3^+ group present in the zwitterionic amino acid molecules. Out of 52 crystal structures examined only three patterns of bonding were found (Figure 1.3). The distribution was as follows: 10 for class I, 25 for class II, 14 for class III and none for class IV. From these results it is clear that the preferred pattern is the three-centred hydrogen bond. This was also confirmed by Taylor *et al.* (1984c), who reported that positively charged N^+-H groups are more likely to form three-centre bonds than uncharged N-H groups, and that intermolecular three-centred bonds are unlikely to occur unless the relatively long H…A interaction is stabilised by an appreciable net positive charge on the hydrogen atom.

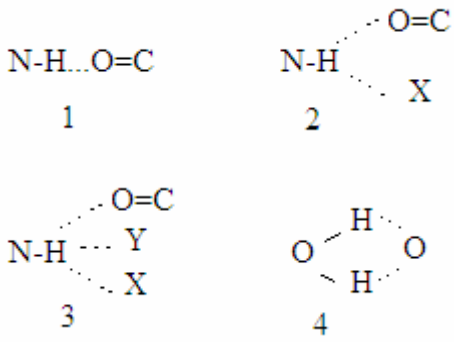


Figure 1.2: Possible hydrogen bond schemes present in a crystal structure.

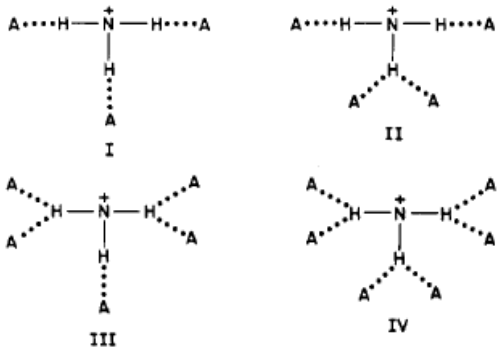


Figure 1.3: Classes of hydrogen bond patterns found in a survey of 52 amino acid crystal structures (Jeffrey *et al.*, 1984).

1.3.5 Non-additivity

In a crystal structure, the energy of an array of *n* interconnected hydrogen bonds is larger than the sum of *n* isolated bonds (Desiraju *et al.*, 1999). This non-additive property arises from the ability of the donor and acceptor groups to form hydrogen bonds. There are two main mechanisms responsible for these non-additive effects.

1.3.5.1 σ-bond cooperativity

In a $D^{\delta-}$-$H^{\delta+}$…$A^{\delta-}$ hydrogen bond, the donor atom becomes more polar by donating the H atom to the acceptor, and this polarity can be increased if the donor atom accepts a H atom itself ($H^{\delta+}$…$D^{\delta-}$-$H^{\delta+}$…$A^{\delta-}$). Thus, functional groups acting simultaneously as hydrogen bond donors and acceptors form infinite chains or extended rings in which the individual hydrogen bonds improve each other's strengths by mutual polarisation (Desiraju *et al.*, 1999). This effect is often called "σ-bond cooperativity" due to the flow of charges along the D-H σ-bonds. An example of a molecular system where this effect takes place is urea (Morrison *et al.*, 2003), which presents an average hydrogen bond energy of 4.9 kcal mol^{-1} (20.3 kJ mol^{-1}) in a periodic system (i.e. crystal environment), and 3.6 kcal mol^{-1} (15 kJ mol^{-1}) in an

aperiodic system (i.e. isolated environment). Thus, the effect of the σ-bond cooperativity calculated on the hydrogen bond energy in urea is of 26%, making the hydrogen bonds stronger and more energetic. Steiner (2002) also reported that the contribution from the σ-bond cooperativity to the strength of intermediate hydrogen bonds is of the order of 20% relative to isolated interactions.

The σ-bond cooperativity effect will be further investigated during the work presented in this thesis, especially on the crystal structures of amino acids.

1.3.5.2 Resonance-Assisted Hydrogen Bonding (RAHB)

The concept of *resonance-assisted hydrogen bonding* (RAHB) was firstly introduced by Gilli *et al.* in 1989, who reported that "the interplay between hydrogen bond and heterodynes (or more generally heteroconjugated systems) can strengthen remarkably the hydrogen bond itself". More recent descriptions of this phenomenon have been given by Gilli *et al.* (1994), there is a further class of strong or very strong hydrogen bonds which cannot be accounted for by electric charges or steric hindrance, but is due to the fact that the neutral donor and acceptor atoms are connected by a system of π-conjugated double bonds", and Desiraju *et al.* (1999) who wrote: *charge flow in suitably polarisable π-bond systems increases donor and acceptor strengths*. Many studies have been carried out to obtain a better understanding of the magnitude of this effect (Alkorta *et al.*, 2004; Mohajeri, 2004).

1.4 Polymorphism

1.4.1 Definition

The word "polymorphism" is derived from the Greek words *poly* and *morph*, which mean "many" and "form" respectively. Thus, polymorphism literally means "many forms". The first use of the word, in the context of crystallography, was made

by Mitscherlich (1822, 1823), who recognised different crystal structures of the same compound in a number of arsenate and phosphate salts. Nevertheless, the simplest definition of "polymorphism" was given by Rosenstein *et al.* (1969): *when a substance can exist in more than one crystalline state it is said to exhibit polymorphism*. This definition was later adopted and modified and it is still in constant development.

1.4.2 Polymorphism and the hydrogen bond

A crystal structure corresponds to a free energy minimum that is not necessarily the global minimum (Bernstein, 2002). Moreover, the existence of a similar energy spaced local minimum is the thermodynamic reason for the possible formation of polymorphism, assuming that these minima are kinetically accessible. Molecules adopt different environments in the various polymorphs of a compound. This means that the intermolecular interactions acting between the molecules might not be the same (Figure 1.4). The hydrogen bonding present in a crystal structure can be related to the sublimation of the crystal. Thus, the separation of a molecule from its crystal environment needs to overcome all the attractive forces acting on it, and the energy involved in this process is the so-called *sublimation energy*. The energetic justification for the different molecule conformations in different polymorphs is that the differences in lattice energy among different polymorphic forms can be expected to be around 1 or 2 kcal mol^{-1} (0.25-0.5 kJ mol^{-1}) (Bernstein, 2002). This difference in energy is mainly due to changes in the torsional angles around the single bonds, rather than distortions in bond distances and angles, which are more energetic. This change in the torsional angles enables the molecules present in a crystal structure to modify their conformation, giving rise to *conformational polymorphism,* which is the existence of different molecule conformations in different polymorphs.

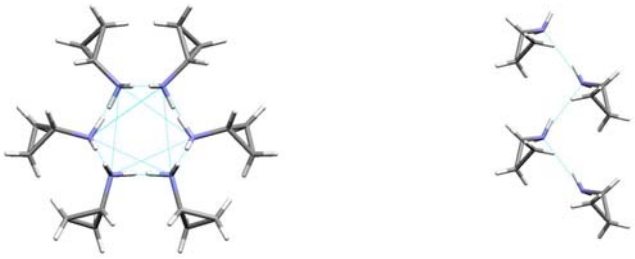


Figure 1.4: Polymorphism in cyclopropylamine: Phase I (left), Phase II (right).

1.4.3 Techniques to study polymorphism

The phenomenon of Polymorphism can be investigated through many different techniques, such as X-ray diffraction, solid-state nuclear magnetic resonance (SSNMR) and infra-red and Raman spectroscopies. Among these, whether it is single crystal or powder, X-ray diffraction constitutes a definitive method for the determination of molecular structure and it is the principal method used to produce the experimental work presented in this thesis. The experimental techniques will be described in detailed in Chapter 2, together with the different computational methods used in combination with the experiments, to fully describe the crystal structure of molecular systems.

1.4.4 Importance of Polymorphism

Polymorphism has become of great interest in many research areas, from Inorganic chemistry, where the focus of interest is on the different properties (electronic, magnetic, etc.) of solids, to Organic and Medicinal chemistry, which studies the use of different polymorphic forms to target different diseases, to improve toxicity and solubility problems and minimise the production cost. Examples of this are paracetamol (Fabbiani *et al.*, 2003) and aspartame. The patent case for the latter was described by Bernstein in his book “Polymorphism in molecular crystals” (2002).

1.5 Description of hydrogen bonded systems

The description of hydrogen bonded molecular systems can be made in terms of their topology. The methods used to produce the work presented here are graph set notation, Hirshfeld surfaces and Voronoi-Dirichlet polyhedra. Each of these techniques is distinct. However, they all give a description of the bonding and packing behaviour and allow us to classify and visualise the crystal structures.

1.5.1 Graph Set Analysis

The hydrogen bond network of a crystal structure can be "decoded" into different hydrogen bond motifs. This can be done with the use of graph sets, which were introduced by Wells (1962) and used later on by Kuleshova *et al.* (1980) among others, simplifying significantly the description of hydrogen bonded systems. The graph sets used during the production of this thesis describe the hydrogen bond network in terms of number of donors and acceptors present in the pattern, and the nature of the pattern (Etter *et al.*, 1990; Grell *et al.*, 1999). In this way, the description of the hydrogen bonding does not depend on the geometry of the interactions, but on their topology.

A graph set is specified using the pattern designator (G), its degree (r), and the number of donors (d) and acceptors (a), as shown:

$$G_d^a(r)\,.$$

The designator, G, specifies the type of pattern formed by the hydrogen bonds. There are four simple patterns: chains (C), rings (R), intramolecular (S) and other finite patterns (D). The degree of the pattern, r, is the number of bonds present in a chain or the number of atoms present in a ring.

The hydrogen network can be divided into motifs and levels. A *motif* is a pattern containing only one type of hydrogen bond, whereas a *level* is the combination of

various motifs. If the motifs contain only one type of hydrogen bond, the level formed will be *unitary* (first level), whereas if there are two or three different interactions, the level will be *binary* (second level) or *ternary* (level three).

The graph set notation has been used to study the hydrogen bonding in different polymorphic systems, in order to find differences and similarities between them (Bernstein, 1991) and will therefore be used in the following chapters to describe the hydrogen bonding in different biological and organic molecular systems.

1.5.2 Hirshfeld surfaces

The use of Hirshfeld surfaces (Hirshfeld, 1977) makes the investigation of packing modes and intermolecular interactions in molecular crystals possible (McKinnon *et al.*, 2004). The Hirshfeld surfaces divide the crystal into regions where the electron distribution of a sum of spherical atoms for the molecule (the promolecule) dominates the corresponding sum over the crystal (the procrystal). The full computational procedure used to construct the Hirshfeld surface was explained in detail by McKinnon *et al.* (2004) and therefore, only the information necessary to understand the Hirshfeld surfaces will be given in this section. Thus, a weighting function $w(r)$ can be defined for a particular molecule as follows:

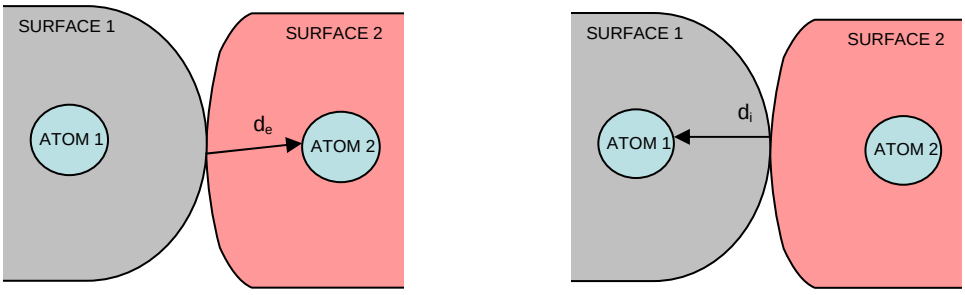
$$\begin{aligned} w(r) &= \sum_{a \in molecule} \rho_a(r) \Big/ \sum_{a \in crystal} \rho \sum_a \rho_a(r) \\ &= \rho_{promolecule}(r) / \rho_{procrystal}(r) \\ &\cong \rho_{molecule}(r) / \rho_{crystal}(r) \end{aligned} \tag{1.1}$$

where $\rho_a(r)$ is a spherically averaged Hartree-Fock atomic electron density function centred on nucleus *a*, and the ratio between promolecule and procrystal electron densities can be regarded as an approximation to the ratio between true molecule and crystal electron densities.

This equation shows that the volume in which the promolecule dominates the procrystal electron density is the region where $w(r) \geq 0.5$. Therefore, the Hirshfeld surface is defined by $w(r) = 0.5$.

Hirshfeld surfaces are not only dependent on the molecular geometry; they are defined inside the crystal and therefore, they reflect the interplay between different atomic sizes and intermolecular contacts in the crystal (i.e. intermolecular interactions). Hirshfeld surfaces pack very close in the crystal but they can never overlap. However, they can get close enough to touch, although this would be an extreme case. Finally, the surfaces leave small intermolecular voids, which can be seen as zones where the crystalline electron density is very low and is not dominated by any single molecule.

In addition to the Hirshfeld surfaces, two-*dimensional fingerprint plots* can be used to study intermolecular contacts in crystal structures. In order to understand the meaning of these fingerprint plots, the external distance (d_e) and internal distance (d_i) to the surface must be defined. The *distance external to the surface*, d_e, measures the distance from the surface to the nearest nucleus in another molecule [Figure 1.5(a)], whereas the *distance internal to the surface*, d_i, measures the distance from the surface to the nearest atom in the molecule [Figure 1.5(b)].



(a) (b)

Figure 1.5: Schematic representation of two Hirshfeld surfaces constructed for two different molecules. We can define (a) d_e as the external distance from the surface to the nearest atom in another molecule (i.e. from surface 1 to atom 2) and (b) d_i as the internal distance from the surface to the nearest atom in the molecule itself (i.e. from surface 1 to atom 1).

The two-dimensional fingerplots are built by plotting (d_i, d_e) pairs. This means that the plots will be mostly symmetrical. However, due to the presence of voids in the structure the symmetry might be altered. Nevertheless, the fingerprint plots will keep a high degree of symmetry (Figure 1.6). These fingerprint plots, derived from the Hirshfeld surface, summarise the frequency of each (d_i, d_e) combination across the surface of a molecule, indicating not only which interactions are present, but also the relative area of the surface corresponding to each such interaction. The colour of the areas in the fingerprint plots tell us about the contribution from the surface triangles (McKinnon *et al.*, 2004). Typical features of these fingerprint plots are: the formation of sharp spikes, which are characteristic of strong hydrogen bonds, the presence of spikes where $d_e \cong d_i$ is typical of H…H interactions, and the presence of red coloured areas in the centre of the plot, characteristic of π-π stacking interactions.

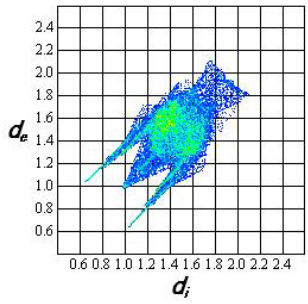


Figure 1.6: Two-dimensional fingerprint plot for the structure of monoethanolamine at 1.6 GPa.

Going back to the description of the Hirshfeld surfaces, some of their properties, such as curvedness and shape index, can be useful to fully understand their meaning. The *curvedness* of a Hirshfeld surface is a measurement of how much shape the surface exhibits [Figure 1.7(a)]. Flat areas of the surface have a low curvedness, whereas areas of sharp curvature have a high curvedness. This curvedness can be used to define a coordination number in the crystal, since areas of the surface with high curvedness tend to divide the surface into contact patches with each neighbouring molecule. Similarly, the *shape index* measures the shape of the surface [Figure 1.7(b)], which gives information about how molecules pack in the crystal. If the Hirshfeld surface exhibits flat faces on regions pointing at other atoms or molecules, the

coordination number of the molecules can be known by counting the number of flat faces. In addition to this, the shape of the surface is very useful to study polymorphism, since the Hirshfeld surfaces for the same molecule can be compared in different crystal structures. The shape index can be sensitive to very subtle changes in the surface shape, especially in regions where the total curvedness is very low. If two shapes are different only in the sign, they will be complementary, “stamp” and “mould” pairs, meaning that maps of shape index on the surface can be used to identify complementary hollows (shape index <1) and bumps (shape index >1).

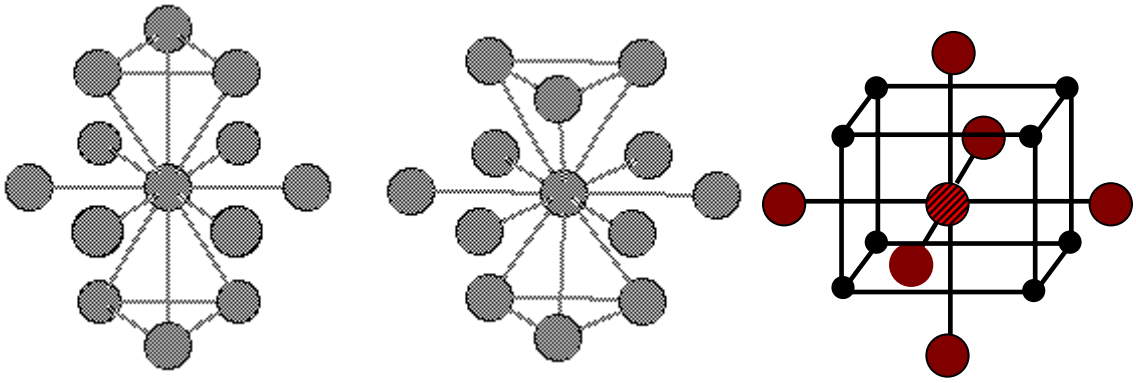
(a) (b)

Figure 1.7: Hirshfeld surfaces for an monoethanolamine molecule in the crystal environment at 1.6 GPa. Each molecule is shown with the Hirshfeld surface mapped with (a) curvedness, and (b) shape index.

1.5.3 Voronoi-Dirichlet Polyhedra

There are many ways to describe molecular systems, some of which have already been explained (graph set notation and Hirshfeld surfaces). The most traditional approach is based on the related close-packing model of hard spheres. The most efficient packings are known to be hexagonal (HCP), and cubic close packing (CCP), in which each sphere has a coordination number of 12 (Figure 1.8). In the slightly less efficient body centred cubic (BCC) each sphere is coordinated by 14 others. The close packing model description is based on representing structural groups, such as molecules, by spheres with different deformability. This is different from the so-called “close-packing model” (Kitaigorodskii, 1973), in which molecules are represented by solids of fixed shape. Filling the whole empty space and mutually

deforming the soft spheres, ultimately generates convex polyhedra (Voronoi-Dirichlet polyhedra, VDP's) thus forming a normal (face-to-face) partition of space (Peresypkina *et al.*, 2000a, 2000b).



(a) (b) (c)

Figure 1.8: View of the arrangements of neighbour spheres in: (a) a CCP [3+6+3], (b) HCP [3+6+3], and (c) BCC [8+6] packings. The coordination number of spheres is 12 for CCP and HCP, whereas for BCC is 14.

The topology of a molecular packing can be visualized using VDPs, which are built from a sublattice of molecular centroids. This method allows us to partition space amongst points occupying that space (Figure 1.9). Thus, a point is separated from a neighbouring point by a line bisecting the vector between them. By repeating this operation for every pair of points, a subdivision of the space is made of cells containing one single point each, which is the centroid.

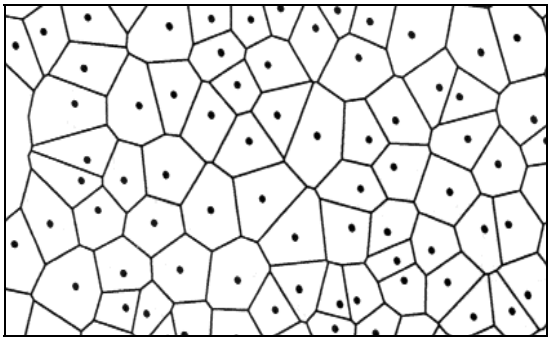


Figure 1.9: View of the partition of a two dimensional space into Voronoi-Dirichlet cells (Byers, 1992).

There are two types of VDPs that can be contructed. The first type is the *molecular VDP*, which is defined as the union of VDPs of the atoms comprising the molecule and can be interpreted as an image of a molecule in a crystal field. The second type is the *lattice VDP*, which is a polyhedron formed by intersecting planes that bisect perpendicularly the lines joining atomic centres (Baburin *et al.*, 2004). The lattice VDP is a simpler method of description as it is formed from the geometrical centroids of the molecules, whereas the molecular VDP takes into account intra and intermolecular interactions. There is a type of molecular VDP, which is called *smoothed molecular VDP* and that is constructed by ignoring the interior structure of contacting molecules, only considering their centroids. This VDP is convex and its shape characterizes the arrangement of molecules around the central one. The number of faces of a smoothed molecular VDP is equal to the molecular coordination number (MCN), which is defined as the number of molecules, which have at least one contact with a given molecule (Kitaigorodskii, 1973). In addition to the MCN, the number of neighbours in the first, second and third coordination spheres can help to characterize the molecular packing. Thus, CCP, HCP and BCC exhibit coordination sequences of 12-42-92, 12-44-96 and 14-50-110.

Thesis outline:

The current chapter was dedicated to the introduction of the hydrogen bond in terms of its nature and properties. Therefore, the hydrogen bond has been defined, decomposed into its constituent interactions and classified depending on its

characteristics. Then some important aspects of the hydrogen bond have been discussed, such as the influence of σ and π-bond cooperativities on its strength and geometry. The second part of the introductory chapter was dedicated to the study of polymorphism, its relationship with hydrogen bonding and the different techniques used to study polymorphic changes. Finally, the description of hydrogen bonded systems in terms of graph set notation, Hirshfeld surfaces and topological analysis has been dealt with.

The second chapter is a concise description of the experimental and computational techniques used during the work presented in this thesis. The high-pressure experimental methodology will be explained in detail in the first part of the chapter, together with the description of the diamond anvil cell used to carry out the experiments. The second half of the chapter will introduce and briefly describe the fundamentals of computational chemistry with particular reference to Density Functional Theory.

Chapter three illustrates the behaviour of cyclopropylamine at ambient and high pressure. A new polymorph of this compound is obtained after crystal growth from liquid at 1.2 GPa (Phase II). As there is either no detailed description in the literature of the known low-temperature (Phase I) polymorph structure or the structure of the polymorph has yet to be determined (Phase II), the crystal structures of the polymorphs at both ambient and high pressure will be described in detail. In addition, Hirshfeld surfaces and topological analyses will be used to explain the molecular environment present in the crystal structure of the different structures. Finally, a computational detailed study is performed on the different phases of cyclopropylamine to study the energetics of the hydrogen bonds present in the crystal structures.

The fourth chapter describes the computational study of α-glycine, which will be used as a test case for the study of the other more complex amino acid systems presented later. This chapter contains details of a new computational method used to study the relationship between the geometry and energy of the hydrogen bonds present in the crystal structure of α-glycine. Once the computational method is proven

to give reasonable structural and thermodynamic data, such as sublimation energies, comparable with those obtained through experimental work, the energies of the individual hydrogen bonds will be calculated in order to extend the work performed on α-glycine to other amino acids which lack experimental data. Additionally, the compressibility of the α-glycine crystal structure will be studied in terms of the energies of the hydrogen bonding.

Chapters five and six describe the experimental and computational work performed on L-α-aspartic acid and L-α-glutamine to study the compressibility of these systems at high pressure, as well as its relationship with the energies of their characteristic hydrogen bonds. The computational method used in the test case of α-glycine will be used during the analysis of the two larger amino acids. No phase changes were observed in any of the two cases presented here up to pressures of approximately 5-6 GPa. Finally, the environment of the amino acid molecules will be studied in terms of Hirshfeld surfaces and topological analyses.

Finally, chapter seven will give some conclusions from the work presented during this thesis on together, biological and small organic systems and will suggest future work avenues for this line of research.

1.6 References

Alkorta, I.; Elguero, J.; Mó, O.; Yáñez, M.; Del Bene, J. E. (2004). *Mol. Phys.* **102**, 2563-2574.

Baburin, I. A.; Blatov, V. A. (2004). *Acta Cryst.* **B60**, 447-452.

Bernstein, J. (1991). *Acta Cryst.* **B47**, 1004-1010.

Bernstein, J. (2002). *Polymorphism in molecular crystals*, Oxford University Press, Oxford.

Brown, I. D. (1976). *Acta Cryst.* **A32**, 24-31.

Byers, J. A. (1992). *J. Animal Ecology* **61**, 759-768.

Del Bene, J. E.; Jordan, M. J. T. (2001). *J. Molec. Str.* **573**, 11-23.

Desiraju, G. R.; Steiner, T. (1999). *The weak hydrogen bond in structural chemistry and biology*, Oxford University Press, Oxford.

Emsley, J. (1981). *Chem. Soc. Rev.* **9**, 91-124.

Etter, M. C.; McDonald, J. C.; Berstein, J. (1990), *Acta Cryst.* **B46**, 256-262.

Fabbianni, F. P. A; Allan, D. R.; Dawson, A.; David, W. I. F.; McGregor, P. A.; Oswald, I. D. H.; Parsons, S.; Pulham, C. R. (2003). *Chem. Comm.* **24**, 3004-3005.

Fortes, A. D.; Brodhott, J. P.; Wood, I. G.; Vocadlo, L. (2003). *J. Chem. Phys.* **118**, 5987-5994.

Gilli, G; Gilli, P. (2000). *J. Mol. Str.* **552**, 1-15.

Gilli, G.; Belluci, F; Ferretti, V.; Bertolasi, V. (1989). *J. Am. Chem. Soc.* **111**, 1023-1028.

Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. (1994). *J. Am. Chem. Soc.* **116**, 909-915.

Grell, J.; Berstein, J.; Tinhofer, G. (1999). *Acta Cryst.* **B55**, 1030-1043.

Hay, B. P.; Dixon, D. A.; Bryan, J. C.; Moyer, B. A. (2002). *J. Am. Chem. Soc.* **124**(2), 182-183.

Hirshfeld, F. L. (1977). *Theor. Chim. Acta* **44**, 129.

Ireta, J.; Neugebauer, J; Scheffler, M. (2004). *J. Phys. Chem. A* **108**, 5692-5698.

Jeffrey, G. A. (1997). *An introduction to hydrogen bonding*, Oxford University Press, New York.

Jeffrey, G. A.; Mitra, J. (1984). *J. Am. Chem. Soc.* **106**, 5546-5553.

Kitaigorodskii, A. I. (1973). *Molecular crystals and molecules*. New York: Academic Press.

Kuleshova, L. N.; Zorky, P. M. (1980). *Acta Cryst.* **B36**, 2113-2115.

Latimer, W. M.; Rodebush, W. H. (1920). *J. Am Chem. Soc.* **42**, 1419-1433.

McKinnon, J. J.; Spackman, M. A.; Mitchell, A. S. (2004). *Acta Cryst.* **B60**, 627-668.

Mitscherlich, E. (1822, 1823). *Abhl. Akad. Berlin*, 43-48.

Mohajeri, A. (2004). *J. Mol. Str.* **678**, 2001-205.

Morokuma, K. (1971). *J. Phys. Chem.* **55**, 1236-1244.

Morokuma, K. (1977). *Acc. Chem. Res.* **10**, 294-300.

Morrison, C. A.; Siddick, M. M.; (2003). *Chem. Eur. J.* **9**, 628-634.

Morrison, C. A.; Siddick, M. M.; (2004). *Angew. Chem. Int. Ed.* **43**, 4780-4782.

Morrison, C. A.; Siddick, M. M.; Camp, P. J.; Wilson, C. C. (2005). *J. Am. Chem. Soc.* **127**, 4042-4048.

Peresypkina, E. V.; Blatov, V. A. (2000a). *Acta Cryst.* **B56**, 501-511.

Peresypkina, E. V.; Blatov, V. A. (2000b). *Acta Cryst.* **B56**, 1035-1045.

Perrin, C. L.; Nielson, J. B. (1997). *Ann. Rev. Phys. Chem.* **48**, 511-544.

Pimentel, G. C.; McClellan, A. L. (1960). *The hydrogen bond*, Freeman, San Francisco.

Rosenstein, S.; Lamy, P. P. (1969). *Am. J. Hospital Pharmacy* **26**, 598-601.

Scheiner, S.; Kar, T.; Gu, Y. (2001). *J. Biol. Chem.* **276**, 9832-9837.

Steiner, T. (2002). *Angew. Chem. Int. Ed.* **41**, 48-76.

Steiner, T.; Kanters, J. A.; Kroon, J. (1996). *Chem. Commun.* **11**, 1277-1278.

Steiner, T.; Saenger, W. (1993). *J. Am. Chem. Soc.* **115**, 4540-4547.

Taylor, R.; Kennard, O. (1984a). *Acc. Chem. Res.* **17**, 320-326.

Taylor, R.; Kennard, O, Versichel, W. (1984b). *Acta Cryst.* **B40**, 280-288.

Taylor, R.; Kennard, O, Versichel, W. (1984c). *J. Am. Chem. Soc.* **106**, 244-248.

Tsuzuki, S.; Luthi, H. P. (2001). *J. Chem. Phys.* **114**, 3949-57.

Wells, A. J. (1962). *Structural Inorganic Chemistry*, Oxford: Clarendon Press.

Chapter 2

Introduction to Experimental and Computational Techniques

2.1 Introduction

After the discovery of X-rays at the end of the nineteenth century, many uses have been given to this type of radiation; from Medicine, to perform X-ray analysis of the human body in order to understand its nature and diagnose illnesses, to Chemistry, Physiscs and Geochemistry, to perform structural analysis of molecular and inorganic materials, as well as conductor and semiconductor compounds. The use of X-rays in Chemistry became especially important once it was realised that they would be useful for probing the three-dimensional structure, at an atomic level, of crystalline materials. As the wavelength of X-rays is of the same order of magnitude than the spacings, or distances, between atoms, an incident X-ray beam could be made to diffract from a single crystal, giving rise to a pattern of diffraction spots, which can be recorded on photographic film and used to deduce the arrangement of the atoms in the crystal

The first X-ray diffraction studies carried out to investigate the solid state structure of molecular materials were based on extremely simple structural systems, such as rock salt (Bragg & Bragg, 1915). However, over the last few decades, as experimental and data analysis techniques advanced, much more complex and challenging systems, such as penicillin (Crowfoot *et al.*, 1949) and vitamin B12 (Hodgkin, 1955), could be addressed. Nevertheless, the structural studies of these complex systems remained far from routine experiments. Furthermore, with the development of fully automated diffractometers, improvements in detector technology and advances in X-ray sources, from in-house rotating anode generator to third-generation synchrotron sources, very much more complex systems, such as viruses (Klug et al., 1957; Steere et al., 1958; Sukihara, 1990), proteins (Bragg, 1953) and even DNA fragments (Drew, 1981; Lipps *et al.*, 2004) can have their structures determined down to the atomic level. In addition to this, and in parallel to these advances in the experimental techniques, a rapid improvement in data analysis and visualisation software has taken place, partially due to the necessity to cope with the very large volumes of data required to determine these very large structures. Despite all of these advances, however, the expertise and intuition of the crystallographer remains crucial to the successful structure solution of more difficult problems.

The use of *ab initio* computer simulations have been of great use to complement X-ray diffraction studies so that a more complete characterisation of a material can be achieved. Although *ab initio* calculations have been extensively carried out on systems in the gas phase due to the simplicity of the systems investigated, such as the study of ammonia (Tinland, 1968; Sadle *et al.*, 1986) and water (Mantz *et al.*, 2005; Burnham *et al.*, 2002) the development of new algorithms and methods have made possible the use of *ab initio* calculations to study the structural features of more complex crystalline systems –inorganic and organic–, in the solid state (Bernasconi *et al.*, 1995; Fortes *et al.*, 2003). Advances in software, such as Vasp, Castep and Crystal, which implement density functional formalism and delocalised/localised basis sets, have made possible the study of inorganic and organic materials in the solid state, giving reliable results which, in general, are in good agreement with the experimental results. Examples of the applications of *ab initio* computer simulations are the geometry optimisation of the structure to locate hydrogen atoms accurately (Allan *et al.*, 1999), which can be a difficult task when using X-rays, and to determine the relative energies and stabilities of polymorphs (Chisholm *et al.*, 2005), as well as determine the strengths of individual bonds.

During the production of this thesis, X-ray diffraction techniques and *ab initio* computer simulations were used in combination, in order to get a better description of the crystalline materials studied, but also to prove, once more, that the partnership between both techniques allows the achievement of a better picture of the crystal structure, giving accurate information about the structural features (through experimental methods) present in the system, but also the thermodynamic properties (through computational methods).

In the next two sections, the experimental and computational techniques used to produce the work here presented in this thesis will be described in detail.

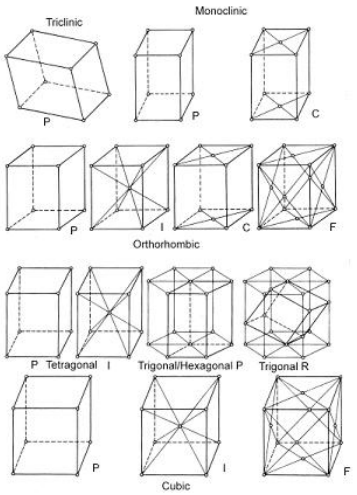
2.2 Experimental techniques

2.2.1 Introduction

In late 1895, a German physicist, W. C. Roentgen (Nobel prize, 1901) realised that he had produced a previously unknown "invisible light", or ray, which he later named the X-ray, as "X" is used in mathematics to indicate the unknown quantity. Following the discovery of X-rays, the X-ray diffraction theory was developed, first by Max von Laue (Nobel prize, 1914) and Paul Knipping, who proved that X-rays were not particles, but waves of light with very small wavelengths and later, it would be William Lawrence Bragg who would advance the theory to allow full structures to be determined. Bragg confirmed what was already known, that a crystal is arranged as a lattice, suggesting that the distance between molecules in a solid is around a tenth of a nanometre. The term lattice refers to an infinite array of discrete points with an arrangement and orientation that appears exactly the same when viewed from any point of the array. The regular arrangement of atoms in a crystal constitutes a lattice. In 1848, Auguste Bravais demonstrated that in a three dimensional system there are fourteen possible lattices, which can be grouped into seven systems, as shown in Table 2.1.

Among the analytical and spectroscopic techniques, such as Infrared (IR) and Raman spectroscopy, Nuclear Magnetic Spectroscopy (NMR), Mass Spectroscopy and Elemental Analysis, X-ray diffraction is the only technique which gives direct information about the structural arrangements of atoms and molecules in the solid state. As X-rays are scattered by electrons, the scattering power of an atom of a particular element is related to its atomic number. This makes light atoms, particularly hydrogen, difficult to locate. This problem can be resolved with the use of neutron diffraction. Neutrons can also be diffracted by crystals and their scattering power is almost random with regard to atomic number. This allows the location of the nuclei of the different atomic constituents of the molecules present in the structure. As hydrogen scatters neutrons incoherently giving rise to a noisy background, which complicates significantly the data processing, neutron diffraction constitutes a better

tool to look at hydrogen bonded systems, which becomes especially important when locating hydrogen atom positions at high pressure. Techniques like Atomic Force Microscopy (AFM) or Scanning Tunnelling Microscopy (STM), which provide images of atoms, or ions in surfaces with near atomic resolution, work particularly well for isolated molecules but fail for more volatile systems. One technique that can be used to study volatile compounds, both inorganic and organic, is gas-phase electron diffraction because electrons have both wave and particle properties, so they are diffracted by atoms and molecules, but can also be observed as distinct particles. In summary, X-ray and Neutron diffraction techniques can be used in combination with the aforementioned techniques in order to study the structure features of many types of materials, whether they are in the gas, liquid or solid phase.



System	Axials lengths and angles	Bravais lattice	Lattice Symbol
Cubic	Three equal axes at right angles $a=b=c$, α=β=γ=90°	Simple Body-centered Face-centered	P I F
Tetragonal	Three axes at right angles, two equal $a=b\neq c$, α=β=γ=90°	Simple Body-centered	P I
Orthorhombic	Three equal axes at right angles $a=b=c$, α=β=γ=90°	Simple Body-centered Base-centered Face-centered	P I C F
Trigonal (Rhombohedral)	Three unequal faces, equally inclined $a\neq b\neq c$, α=β=γ=90°	Simple	P
Hexagonal	Three unequal axes at 120°, third axis at right angle $a\neq b\neq c$, α=β=90°, γ=120°	Simple	P
Monoclinic	Three unequal axes, one pair not at right angles $a\neq b\neq c$, α=β=γ=90°	Simple Base-centered	P C
Triclinic	Three unequal axes, unequally inclined and none at right angles $a\neq b\neq c$, α≠β≠γ≠90°	Simple	P

Table 2.1: Crystal systems and Bravais lattices (http://www.cmmp.ucl.ac.uk).

Almost a century after the discovery of X-rays, the use of X-ray crystallography is now used routinely in structural research due to the development of new experimental and data analysis techniques. For example, in the late 1960's, the original wet-film technique was replaced by computer automated diffractometers equipped with point detectors. More recently, however, area detectors have come back into favour, initially through image-plate technology and laterly through large CCD area detectors, which were introduced in 1994. These are replacing conventional four-circle diffractometers as the workhorses of single-crystal data collection. The advantage of this is the ability to record more diffraction data in a shorter time with no need for obtaining any information *a priori* about the crystal structure.

Additionally, advances in the X-ray sources have allowed the use of alternative ways of creating X-rays. The sealed X-ray tube is the conventional source of X-rays in most laboratories, which, depending on the metal used as anode, emits monochromatic radiation of defined wavelengths. These wavelengths are chosen according both to the cell parameters of the compound under study and to the chemical elements present. For example, complex systems with large unit cells composed of weakly scattering elements, such as carbon, oxygen and nitrogen, require X-rays produced by copper anodes due to the longer radiation wavelength they produce, whereas simpler systems with high X-ray absorption may require shorter wavelengths from molybdenum or silver anodes. The wavelengths for the most commonly used anodes in X-ray crystallography are molybdenum (0.7135 Å), copper (1.5443 Å) and chromium (2.2935 Å). These defined wavelengths are one of the limitations of the X-ray tubes, and together with the low intensity and high difference of the X-ray beams they produce, has led the increased use of this type of radiation for more challenging problems. The first generation of synchrotron sources were high energy physics accelerators, where the synchrotron radiation was an unwanted by-product. It was soon realised that this radiation could be used to probe the structure of matter and initially generators were built on these electron accelerators to produce a "parasitic" source of X-rays, which not only had a high intensity but covered a wide range of wavelengths. With the success of these initial tests, a second generation of dedicated machines were built –the first of which was the SRS Daresbury laboratory in the UK, which opened in 1980-, and from advances in

accelerator technology a third generation of synchrotron have been built producing a step-change in flux and photon energies. The first of this new breed of synchrotrons, the European Synchrotron Radiation Facility (ESRF), was opened in Grenoble, France, in 1994. Synchrotron radiation facilities are located all over the world (APS, USA; Spring-B, Japan). Thirty years after the construction of the first dedicated synchrotron radiation facility, the Diamond synchrotron (Oxfordshire, UK) is being built and it is expected to become operational in 2007. The use of synchrotron radiation has revolutionised structural science, not only allowing the structures of extremely complex proteins and viruses to be determined at the atomic level, but also allowing studies of materials at extreme temperatures and pressures as well as on ultra-short timescales.

2.2.2 X-ray diffraction theory

Crystals, by analogy with the diffraction of light by an optical grating, should be capable of diffracting radiation that has a wavelength of the same order of magnitude as the interatomic separation; i.e. a wavelength of approximately 1 Å. Bragg's law will be explained in the following section, together with the more rigorous Ewald construction, in order to summarise the principles underlying structure determination by X-ray diffraction.

2.2.2.1 Bragg's Law

W. L. Bragg developed his theories on X-ray diffraction during his initial studies on the crystal structure of sodium chloride during 1912. Bragg's Law relates the diffracted X-ray beam with the so called "Miller planes" within the unit cell. The unit cell is the smallest repeating unit that can generate the entire crystal structure with only translation operations, while the Miller planes are the series of planes which partition the crystal structure (defined by the indices h, k, l), while maintaining the periodicity of the unit cell translations. He first established the geometrical conditions which must be satisfied for the diffraction condition within a crystal. Bragg

considered a coherent, monochromatic and parallel X-ray beam to be incident on the crystal (Figure 2.1).

To derive Bragg's law we can trace two parallel X-ray beams (Ray 1 and Ray 2) which will interact with the planes in the crystal. The spacing between the Miller planes is *d*. Ray 1 reflects off Plane 1 at an angle θ which is equal to its angle of incidence. Similarly, Ray 2 reflects off plane B at the same angle θ. However, Ray 2 must travel a distance 2a further than Ray 1. To satisfy the diffraction condition, distance 2a must be equal to an integer number of wavelengths, nλ, so that Rays 1 and 2 will combine to produce constructive interference. In terms of the beams labelled Ray 1 and Ray 2 in Figure 2.1 this requires that the distance AB + BC be equal to nλ. Accordingly:

$$AB + BC = n\lambda \qquad (n = 1, 2, 3, \ldots) \qquad (2.1)$$

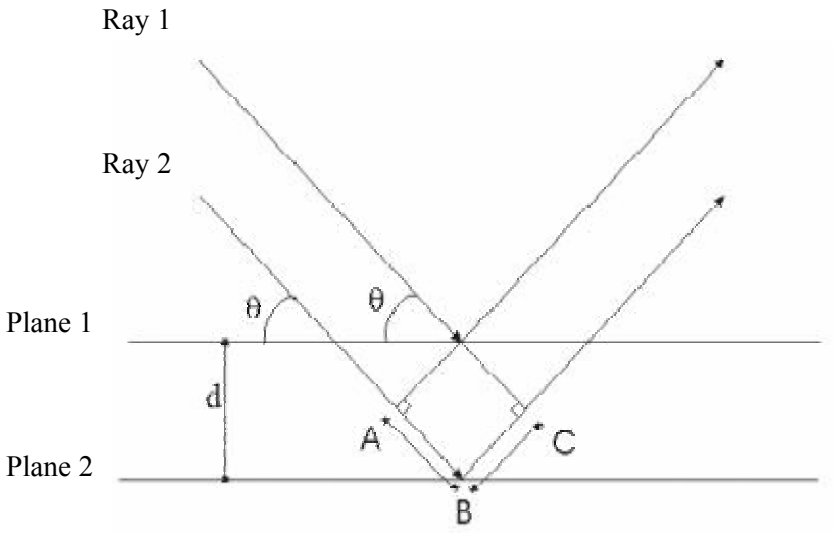


Figure 2.1: Bragg's law, assuming that the planes within the crystal act as reflecting planes (http://adelle.biochem.queensu.ca).

Since AB = BC and $\sin\theta = \dfrac{AB}{d_{hkl}}$ [AB = $d_{hkl} \sin\theta$]:

$$n\lambda = 2d_{hkl} \sin\theta \qquad (2.2)$$

This relation is known as Bragg's Law and describes the angula

r position of the diffracted beam in terms of λ and d_{hkl}. In most experiments, we deal with first order diffraction (n = 1) so the equation (2) can be simplified to:

$$\lambda = 2d_{hkl} \sin\theta \qquad (2.3)$$

2.2.2.2 The Ewald construction

The Ewald construction provides a geometrical framework to link Bragg's Law with the lattice of the unit cell. If the incident and diffracted beams can be considered to be vectors with lengths proportional to the reciprocal of the wavelength, then as θ is varied the vectors will describe the surface of a sphere, the Ewald sphere, with radius $1/\lambda$ (Figure 2.2).

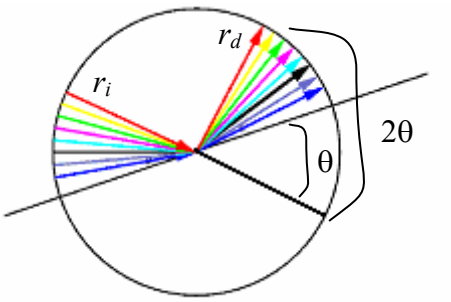


Figure 2.2: View of the circle formed by the incoming (r_i) and outgoing (r_d) X-ray beams when the angle of diffraction θ, is modified.

The angle between the incoming and outgoing beams is always two times θ as is clear from Figure 2.3. It can also be seen the existence of a red vector connecting the points where the diffracted beam, r_d, and the incoming beam, r_i, "exit" the circle,

so S = r_i-r_d. The length of this vector depends on the radius of the circle *R* and the angle θ. Additionally, the vector is perpendicular to the lattice plane for any value of θ and the lattice plane always cuts the vector in half. Thus, half of the vector can be calculated as:

$$S/2 = R\sin\theta \tag{2.4}$$

Therefore,

$$S = 2R\sin\theta \tag{2.5}$$

If R = 1/λ, then

$$S = 2\frac{\sin\theta}{\lambda} \tag{2.6}$$

If we compare this to Bragg's Law, which was derived in equation 2.3 and rearranged for comparison purposes:

$$1/d_{hkl} = 2\frac{\sin\theta}{\lambda} \tag{2.7}$$

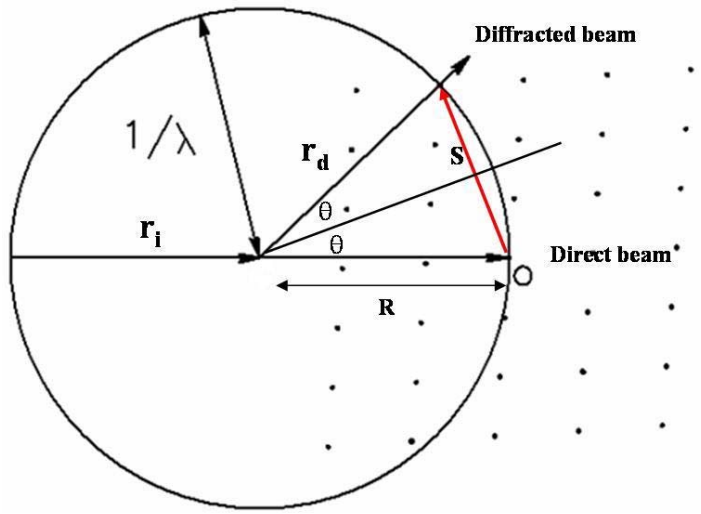


Figure 2.3: The Ewald sphere construction and reciprocal lattice (http://169.232.130.50/khan/nutshell_files).

It can be seen that the length of the vector in this geometric construction corresponds to *1/d* in diffraction theory, where *d* is the lattice spacing. Thus, this vector corresponds to the reciprocal space diffraction vector, S_{hkl}.

The real space origin is defined by the reciprocal lattice, i.e. the point where the horizontal beam exits the circle will be the origin of the reciprocal space, *O* (0, 0, 0). The Ewald construction shows that for any reciprocal space vector *S*, only those that fall on the surface of the sphere will obey Bragg's Law and will give rise to diffraction.

In summary, the Ewald construction is a generalisation of Bragg's Law and allows us to define the three-dimenstional nature of single-crystal diffraction in terms of the reciprocal lattice.

2.2.2.3 Structure Factor Equation

From the experiment a series of reflections are harvested, which each have an exact three-dimensional location within the diffraction pattern and an intensity. The location of the reflections in space, usually in terms of diffractometer angles, give information on the unit cell of the crystal (*via* the reciprocal lattice) while the diffracted intensities are related to the location of the atoms within the unit cell and to the scattering power of the atoms. The intensity of a reflection is related to the underlying crystal structure via the structure factor equation.

The structure factor is a way of expressing a group of atoms in the unit cell as planar elements, developing the diffraction intensities from each of those elements and integrating the results into the total diffraction intensity from each d_{hkl} plane in the crystal structure. The structure factor equation can be written as:

$$F_{hkl} = \sum_{j=1}^{atoms} f_j \exp\left[2\pi i\left(hx(j) + ky(j) + lz(j)\right)\right] \tag{2.8}$$

where f_i are the atomic scattering factors, h, k and l are planes within the reciprocal lattice for each reflection and x, y and z are the positions of the atoms. As the X-rays are actually scattered by the electrons surrounding each atom, the atomic scattering factor scales with the number of electrons that the atom contains.

In Equation 2.8 the atoms are approximated as being vanishingly small points in space. However, as the electron clouds surrounding the atoms have approximately the same diameter as the X-ray wavelength, this causes an incoherence that develops across the electron cloud, which diminishes the overall scattering power especially at higher angle. Figure 2.4 illustrates this behaviour by plotting the scattering factor versus $\sin\theta/\lambda$. Note that at $\theta = 0$ the value of f is simply the number of electrons present in the atom. The vibration of atoms, which appears as a spreading out of the electron density when averaged over the whole crystal, alters the effective X-ray scattering from each atom still further (note that the accuracy of the fitting is higher at higher angles than those at low angles due to the approximations made to fit the data). This effect can be modelled by multiplying the atomic scattering factor by the isotropic temperature factor, $exp(-Bsin^2\theta/\lambda^2)$, where B is related to the mean square atomic displacement U by $B = 8\pi^2U$ and has units of $Å^2$. An anisotropic model for vibrations, which requires six displacement parameters in its description, is often used provided that the data quality is sufficient. The overall effect of atomic vibration on the diffraction pattern is to reduce its intensity at relatively large Bragg angles. To reduce this effect, and thereby improve the resolution of the experiment, diffraction data are often collected with the sample held at low temperature.

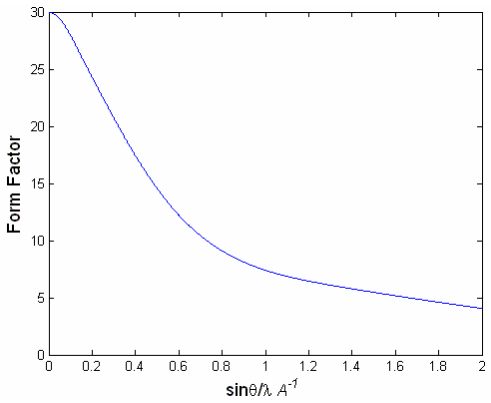


Figure 2.4: View of the fall-off of the atomic scattering f, in units of electrons, with $\sin\theta/\lambda$, this is with the diffraction angle, for Zn (http://tx.technion.ac.il/~katrin).

2.2.3 High-pressure single crystal X-ray diffraction

Despite pressure being one of the fundamental physical constants, pressure-dependent investigations may be more challenging compared with those investigations dependent on temperature due to the necessity, in some cases, of preparing the sample *in situ* under pressure. However, with the development of new simple devices for high-pressure crystallographic studies such as the diamond-anvil cell, over the last half of the 20th century, the mounting and centring of the sample have been drastically simplified in such a way that the data collection does not longer require extra fittings so standard X-ray diffractometers can be used. However, as will be shown in the following sections of this thesis, the use of diamond-anvil cells introduces a unique set of difficulties for data processing.

X-ray diffraction techniques were used in first instance at high pressure by mineralogists, who developed the diamond anvil cell (Merrill *et al.*, 1974; Koepke *et al.*, 1985; Sterer *et al.*, 1990). This high pressure technique was later used by physicists to be applied to many areas of geology and physics, such as the study of periodic elements at very high pressures (Schwarz *et al.*, 1999; Belonoshko *et al.*, 2004a, 2004b). Only more recently have studies been undertaken routinely on molecular systems of interest to chemistry and biology.

2.2.3.1 The diamond anvil cell

The diamond anvil cell (DAC) has become the most widely used device for X-ray diffraction studies under high static pressure, as it is relatively easy to use and the sample volume required for single-crystal studies, to fairly modest pressures, is comparable to that required for a typical ambient-pressure study.

Merrill-Bassett cells (Merrill & Bassett, 1974) were used to carry out all the high-pressure experiments reported in this thesis [Figure 2.5(a)]. The review by Miletich *et al.* (2001) of high-pressure single-crystal techniques offers a detailed explanation of the geometry of a variety of cell designs as well as their components. For the purposes of understanding how the experiments were performed, only the

most general features of the DAC will be described here. The sample is loaded into the pressure cell by positioning it in the pressure chamber created between the flat parallel faces (culets) of two opposed diamond anvils and the hole penetrating a tungsten foil or gasket [Figure 2.5(b)]. A ruby chip, placed beside the sample, is used as a pressure calibrant due to the sensitivity of its fluorescence bands, R_1 and R_2, to changes in pressure (Yen *et al.*, 1992) and because the lines do not change their shape and width appreciably provided that the pressure remains hydrostatic. As the fluorescence signal is strong, the ruby chip needs to be only a few microns in size and hence it does not occupy a significant fraction of the volume within the gasket hole. In addition to this, due to its small size, it does not give rise to appreciable diffracted intensities in the diffraction pattern. The free volume within the pressure chamber is flooded with a pressure-transmitting medium, which in the experiments reported here is a 4:1 mixture of methanol:ethanol. This pressure medium allows a maximum hydrostatic pressure of 10.4 GPa before vitrification occurs. The gasket not only provides an encapsulated chamber within which the pressure transmitting medium and sample is contained but it also provides mechanical support against the enormous shear forces generated at the anvil tips. The tungsten gaskets used during these experiments were initially 250 μm think and were subsequently pre-indented to approximately 100 μm between the anvils prior to loading. Pressure is applied by forcing the diamonds together, which causes the gasket to extrude around the diamond culets, sealing the pressure chamber [Figure 2.5(c)]. It is important to accurately align the culets of the diamonds to attain the highest possible pressures for a given culet size or sample volume. In the Merrill-Bassett diamond anvil cell (Merrill & Bassett, 1974), the diamonds are mounted on two small triangular platens, which are drawn together by three screws. The diamonds are supported by beryllium discs, which serve as X-ray windows and which are fitted into the stainless steel housings. The conical apertures in the steel housings, each of 40° half angle, permit the measurement of 2θ angles up to a maximum of 80° (Figure 2.6).

Diamond anvils are used in the cell due to the fact that diamond is the hardest material known and relatively transparent to electromagnetic radiation over a wide spectral range, from infrared to hard X-rays (5 eV to 10 keV). This allows the use of

the DAC not only during X-ray diffraction experiments but for Raman and Infrared measurements.

There are important limitations arising from the use of the Merril-Basset cell. The diamond anvils and the Beryllium discs cause significant X-ray absorption and background scattering, as can be seen from Figure 2.7. Their conical apertures restrict the reciprocal space accessibility to about 40% (depending on aperture angle) (Figure 2.8), and as a consequence, only an incomplete data set can be measured. Therefore, even when structure solution is possible from such a limited data set, the refinement of the structure and the determination of the temperature factors can prove difficult. The reduced reciprocal space accessibility makes high-pressure studies of low symmetry structures particularly difficult.

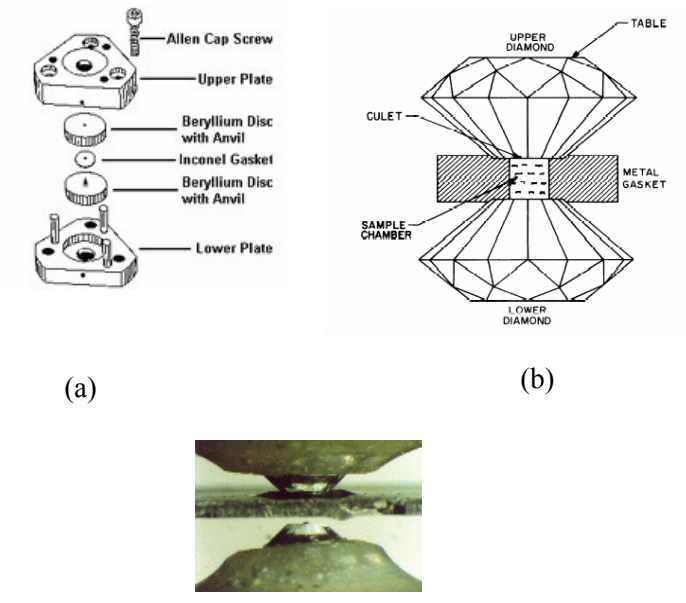


Figure 2.5: (a) A Merrill-Bassett diamond anvil cell with beryllium mounts for X-ray single-crystal diffraction (reproduced from Miletich *et al.*, 2001). (b) Scheme of the diamonds and gasket inside the anvil cell. (c) The diamonds approaching prior to closure of the cell, closing up the sample chamber (http://www.ph.ed.ac.uk/~dra).

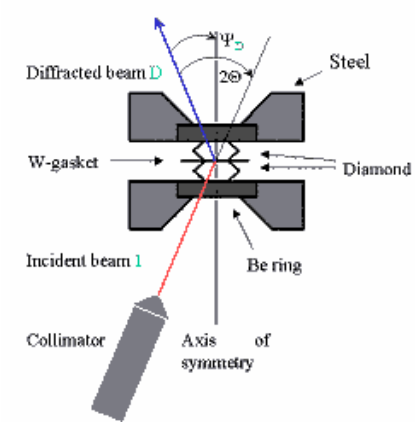


Figure 2.6: The geometry of the incident and diffracted X-ray beams penetrating the diamond anvil cell.

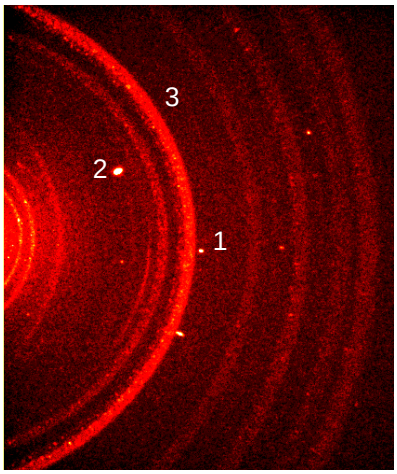


Figure 2.7: Diffraction pattern showing a sample reflection (1), a diamond reflection (2) and the beryllium rings (3).

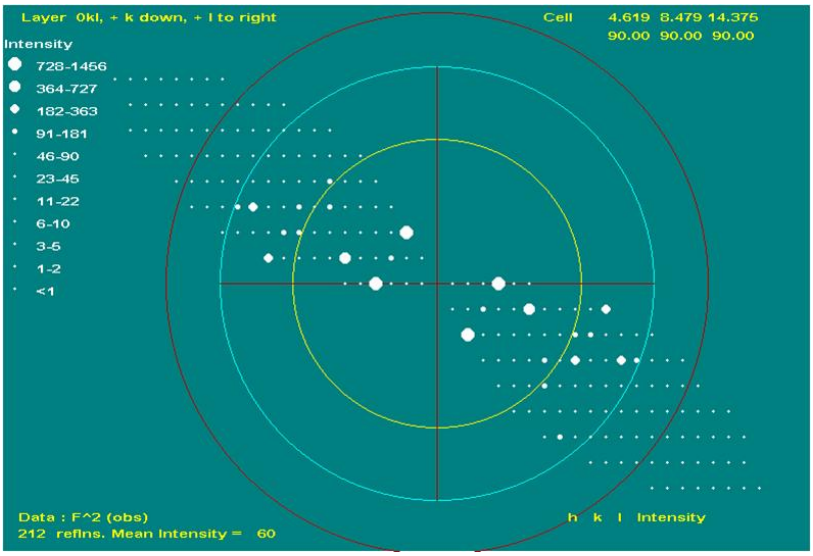


Figure 2.8: Plot of the hkl reflections (including Friedel pairs) after data collection.

2.2.3.2 Preparation of the sample

The preparation of the sample initially involves the crystallisation of the substance and, for high-pressure experiments, this can be achieved using two different methods. The first option is to grow the crystals, from a saturated solution, in the laboratory (*ex situ*) before they are placed in the DAC. For the second choice, the crystals are grown *in situ* from either the melt or saturated solution. Both methods were used in the work presented in this thesis and the detailed methods will be explained in the remaining chapters.

2.2.3.3 Sample centring

Once the DAC is mounted on the goniometer head of the diffractometer, the sample can only be viewed along the direction parallel to the cell axis through the optical parts in the beryllium diamond-anvil supports. The optical centring of the sample along the two directions perpendicular to the crystal is relatively easy to carry out. However, the cell axis must be aligned coaxially with the axis of the alignment microscope and the centring of the sample along this axis can prove difficult. The accurate centring of a sample contained within a DAC was explained in detail by

Dawson *et al.* (2004). The cell is mounted on the goniometer head, assuring that its axis is placed incident to the direct beam (i.e. when $\omega = 0°$) so that it can also be set parallel to the optical axis of the alignment video camera. The cell is then rotated by 180° on ϕ and the camera refocused. The camera is then adjusted to the midpoint of the two micrometer readings on the focus translation and the position of the cell adjusted to re-focus the image of the sample. This procedure is repeated until the sample is in focus with the DAC at 0 and 180° in ϕ.

Finally, the centring of the sample can be checked in the direction of the cell axis as suggested by Meyer *et al.* (2002). The X-ray generator is set up at low power (30 kV and 10 mA), the detector at $2\theta = 0°$ and the sample at $\omega = 25°$ and -25°, and $\phi = 0°$. An image of the beam passing through the gasket hole is taken at these two settings. If the cell is correctly centred the pixel values of the subtracted images at the position of the direct beam should be zero. If this is not the case, small adjustments to the centring can be made to achieve an extremely accurate centring.

2.2.3.4 Data collection

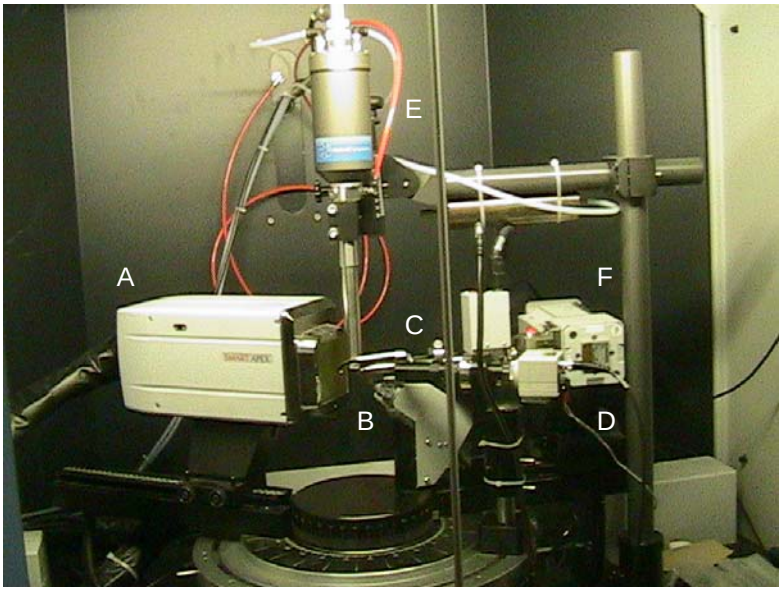
The data collection and processing procedures for the high-pressure experiments presented in this thesis are similar to those described by Dawson *et al.* (2004). The data were collected on the Bruker APEX diffractometer, which is equipped with a charge-coupled detector (CCD) and Molybdenum (Mo-kα) X-ray tube. Figure 2.9 shows the detector, which is the large box on the left side of the instrument. The sample-detector distance is adjustable on the 2θ arm. The distance is usually set at 7 cm for high-pressure experiments. The geometry of the diffractometer circles is illustrated in Figure 2.9. The detector harvests sample reflections from the accessible volume of reciprocal space in a series of eight ω-scans, which are each composed of a sequence of frames in step sizes of 0.3° (Table 2.2). Diffraction images contain background arising from the sample, the diamonds and the beryllium discs, as well as the tungsten gasket, as shown in Figure 2.7. The high background and shading not only produce problems during indexing (due to the spurious reflections generated by the cell components) but also during integration. After integration, an absorption

correction needs to be applied, not only for sample anisotropy but also for absorption by the diamonds and backing disks. Dawson *et al.* (2004) reported some procedures by which these corrections can be applied.

Finally, in cases where the data completeness is very low (e.g. low symmetry crystal systems) the data collection can be performed with the three possible orientations of the DAC mounted on the goniometer head. This "simulates" a rotation in χ.

Run#	2θ/°	ω/°	φ/°	Size/°
1	-28.00	-10.00	90.0	0.3
2	28.00	40.00	90.0	0.3
3	-28.00	-150.50	90.0	0.3
4	28.00	-138.00	90.0	0.3
5	-28.00	-150.50	270.0	0.3
6	28.00	-138.00	270.0	0.3
7	-28.00	-10.00	270.0	0.3
8	28.00	35.00	270.0	0.3

Table 2.2: X-ray diffractometer settings for a typical high-pressure data collection.



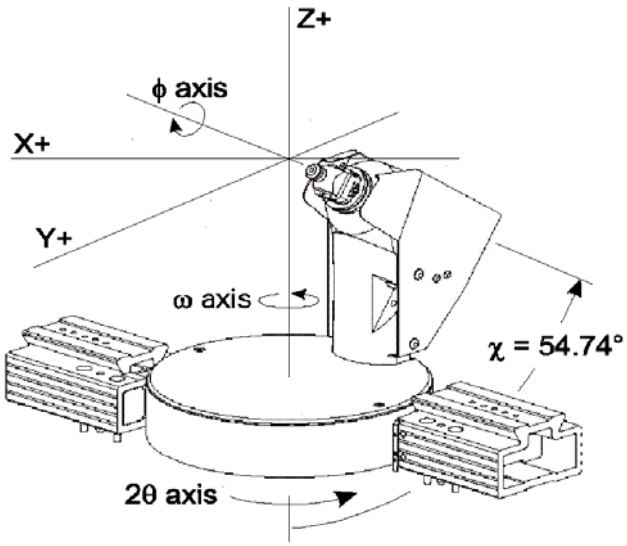


Figure 2.9: View of a Bruker APEX diffractometer used for the diffraction experiments reported in this thesis (reproduced from Bruker APEX diffractometer Manual). This picture shows the charge-couple detector (A), the goniometer head (B) in which the DAC can be mounted, the collimator (C), the camera (D), the cryostream (E) and the X-ray tube (F). The angles and axis defined by the diffractometer components can be found in the scheme at the bottom of the figure.

2.2.3.5 Indexing the data

Once the data have been collected, the first stage of data processing is to index the sample reflections. The diffraction pattern is recorded as a sequence of frames and each frame contains information about the reflections collected in the narrow scan range (usually 0.3°) covered by the frame. The reflections are represented in the form of spots on the CCD image. Their brightness is a function of their intensity and the location of the spot, in terms of pixel position on the frame and the detector and goniometer angles, gives information about the position of the reflection in reciprocal space. Thus, the frame sequence maps out intensity in reciprocal space and by appropriately identifying spots with sample reflections the reciprocal lattice of the crystal can be determined.

For the experiments presented in this thesis, the spots were selected by either picking them by hand or using a thresholding routine in SMART. The reflections within the resulting list were then indexed (identifying the reciprocal lattice) using the program GEMINI (Bruker AXS Inc., 1999). Different solutions can be obtained for the unit cell parameters and the associated orientation matrix UB (which describes the sample orientation with respect to the diffractometer angles). However, only one of these solutions will usually match all the spots satisfactorily and will eventually lead to a full structure solution.

2.2.3.6 Integration

Once the unit cell parameters and the orientation matrix are obtained, the integration of the data (i.e. the determination of the integrated intensity for each reflection) is carried out using SAINT (Bruker AXS Inc., 1997-2001). The integration of the data is initiated by using the unit cell and orientation matrix (UB matrix) obtained from indexing. SAINT will go through all the lattice points predicted by the UB matrix, measuring the intensity of the spots and their position. The program will create an output file containing all integrated reflections and assigning *hkl* indices to all of them. To carry out the integration, the box-size parameters derived from the

observed profile of the peaks in the diffraction pattern are used, which can be constrained during the integration process. The initial orientation matrix obtained from the indexing phase can also be allowed to refine during the initial stages of integration.

One of the major problems in processing the data is the limited volume of the diffraction pattern covered during data collection due to shading by the steel body of the DAC. In conventional experiments, however, SAINT is designed to integrate the entire diffraction pattern which, for the high-pressure data collections, would include spurious reflections from the beryllium and diamonds. If they overlap with predicted sample reflections, SAINT will normally also try to integrate reflections that are predicted to be on shaded regions of the detector. The spurious and shaded reflections would subsequently have to be erased from the datafile before continuing with further data processing. However, to overcome these difficulties masks can be generated for each frame so that pixels shaded by the DAC are not included in the integration. A detailed description about how these "dynamic masks" are generated can be found in Dawson *et al.* (2004).

2.2.3.7 Absorption correction

Due to the different components of the cell (i.e. diamonds, tungsten gasket, backing plates and the bulk of the cell), as well as the presence of the sample, a significant fraction of the incident X-ray beam will be absorbed. The absorption arising from the sample crystal itself is likely to be small. However, the absorption arising from the cell is extremely high and will vary significantly with the path length of both the incident and diffracted beams through the cell components.

The extent of the absorption from the sample depends on the size and shape of the crystal as well as the types and relative proportions of its constituent atoms. The wavelength of radiation used in the experiment is also highly significant. There are three general classes of absorption corrections: analytical (indexing the crystal faces); empirical (collection of more than the minimum of Laue unique data); and Fourier

(solving and refining the data with isotropic displacement parameters). Only the empirical absorption correction was used for the high-pressure experiments described in this thesis. In this method, a comparison is made of the intensity of reflections related by Laue class and from these redundant measurements, an absorption surface for the sample is calculated. This method is typically chosen because it only requires the instrument to collect more data from symmetry related regions of reciprocal space. Area detector data sets usually provide sufficient redundancy for the correction to be applied directly without a significant increase in data collection time. Although there are various data analysis programs that offer this form of absorption correction, SORTAV (Blessing, 1995) was selected for processing the high-pressure data as it is integrated well with the Bruker suite of programs. SORTAV is a multi-purpose program for treating repeated measurements of symmetry equivalent reflections and for subsequently merging the data.

The absorption correction for the DAC was performed by the program Absorb 6.0 (Angel, 2004). A detailed explanation of how the program performs the absorption correction can be found in the Absorb 6.0 manual or in the paper by Angel (2004). This program corrects the data for the absorption arising from the DAC and the gasket shadowing. The Absorb program provides different functions and methods to calculate the DAC absorption correction, which assume that the correction is cylindrically symmetric about the cell axis, and is therefore only a function of the angle φ between the beam direction and the cell axis. It is also assumed that the two halves of the DAC have identical absorption curves.

When X-ray beams enter or leave the DAC at high angles, part of the beams may pass through the gasket, further reducing the measured diffracted intensity. This effect is termed “shadowing by the gasket”. Absorb makes some assumptions when performing the absorption correction for this effect: the gasket hole is a cylinder with an axis parallel to the axis of the DAC, the anvil surfaces are parallel and coincident with the surfaces of the gasket; and finally, all reflections with φ >80° are considered totally obscured by the cell.

2.2.3.8 Structure solution and refinement

Having measured and appropriately corrected the diffraction data, the next step is to solve the structure, where the atomic positions within the unit cell are obtained. If the crystal structure is already known, the coordinates taken from the literature can be used for subsequent refinement. When producing this thesis, some of the atom coordinates were taken from the literature (e.g. L-α-glutamine), whereas in other cases the crystal structure was solved through *direct methods* (e.g. Phase II of cyclopropylamine). The term "direct" is used to describe methods which try to derive the structure factor phases, electron density or atomic coordinates by mathematical means from a single set of intensities taken from the X-ray diffraction experiment.

The programs Crystals (Betteridge *et al.*, 2003) or SHELX (Sheldrick, 1998) are used to do the refinement of the structure by using least squares refinement and by varying the numerical parameters such as fractional coordinates and thermal parameters that describe the structure. This produces the best agreement between the calculated and observed structure factors. The limited volume of accessible reciprocal space can also make anisotropic thermal parameters difficult to refine even when the data to parameter ratio is favourable. The type of refinement (isotropic or anisotropic) depends on the ratio between the number of unique reflections and the number of parameters that are refined. In most of the cases, the structure cannot be refined anisotropically because the number of reflections is not sufficient to do so and the orientation of the crystal in the DAC can play an important factor as the temperature factors directed along the cell axis will be poorly determined. In addition to this, in some cases restraints may need to be applied to the bond distances, angles, thermal and vibrational parameters to get reasonable values. As hydrogen scatters X-rays only very poorly, the measured intensities are relatively insensitive to the hydrogen atom parameters. Therefore, when it is not possible to locate the hydrogen atom positions, they are geometrically constrained to the atoms to which they are bonded.

2.2.4 Low-temperature single crystal X-ray diffraction

2.2.4.1 Preparation of the sample

The sample used is generally a crystal grown *ex situ* in the laboratory, glued to a fibre and mounted on the goniometer. If the sample is a liquid at room temperature then it can be encapsulated in a flame-sealed glass capillary (o.d. 0.5 mm), which will be mounted on the diffractometer. The sample can then be frozen *in situ* with the aid of a cryostream cooler (Oxford Cryostreams) and a crystal can be grown using the laser-assisted zone-refinement procedure of Boese et al. (1994). Both methods were used in this thesis.

2.2.4.2 Data collection

Prior to data collection, the sample must be centred and the temperature set as appropriate. The first step during the data collection will be the collection of a matrix, in order to obtain the approximate unit cell parameters, the orientation matrix and the crystal system. Then, depending on the symmetry of the crystal system, a sphere (triclinic), hemisphere (monoclinic), etc must be collected in order to maximise data completeness and redundancy.

The lack of complicating factors arising from the use of a DAC makes the processing of the data much more straightforward and significantly faster. Additionally, the data completeness, as we would expect, is much higher (normally around 90-100%) and anisotropic refinement is generally possible due to the high data/parameter ratio (usually >10) and the full three-dimensional coverage of the data (Figure 2.10). Due to all these factors, the time required to collect data at low temperature (usually 10 s. per frame, 0.5° scans) is much lower than the time required for the data collection at high-pressure (minimum 20-30 s, 0.3° scans). Finally, due to the quality of the data collected at low temperature, the process of indexing is very straightforward within SMART, and there is no need to manually search for reflections, as is nearly always necessary for high pressure data sets.

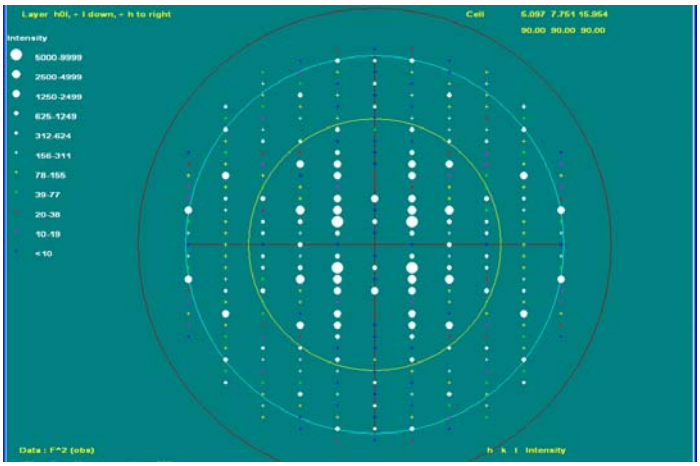


Figure 2.10: Plot of the hkl reflections (including Friedel pairs) after the data collection of L-α-glutamine at 150 K. It can be seen how the amount of reciprocal space collected is much higher than that obtained from collecting data with a DAC.

2.3 Computational techniques

2.3.1 Introduction to quantum mechanical simulations

There are two approaches to computational modelling: *molecular mechanics (MM)*, which is based on classical or Newtonian physics, and *quantum mechanics (QM)*, which is based on de Broglie's wave/particle duality. The first offers the attractive feature of fast computation, which enables very large systems to be studied in reasonable timescales. At the simulation's heart is a force field, which is derived from the transferable properties (geometry and force constants) of functional groups. The simulation is therefore only reliable for systems that are well parameterised by that force field (Boyd & Lipkowitz, 1982). The challenge of building a model to accurately mimic any molecular crystalline system, however, has been met by modern quantum mechanical calculations (Payne *et al.*, 1992). This style of simulation is based on finding an approximate solution to Schödinger's equation. Although most of the work presented in this thesis has been modelled using QM, the use of MM has

been particularly useful to model amino acids in the solid state due to their electrostatic nature, in order to support data obtained through QM.

The fundamental postulate of quantum mechanics (QM) is that a so-called wave function, ψ, exists for any (chemical) system, and that appropriate operators (functions) that act upon ψ return the observable properties of the system. In mathematical notation,

$$\vartheta\psi = e\psi \,, \tag{2.9}$$

where ϑ is an operator and *e* is a scalar value for some property of the system (Cramer, 2003; Hinchliffe, 2001). When equation 2.13 holds, ψ is called an eigenfunction and *e* an eigenvalue. The product of the wave function ψ with its complex conjugate (i.e. |ψ*ψ|) has units of probability density. This notation can be simplified by using real wave functions so that the probability that a chemical system will be found within some region of multi-dimensional space is equal to the integral of $|\psi|^2$ over that region of space. The consequences of these postulates are: (1) the normalised integral $|\psi|^2$ over all space must be unity (i.e., the probability of finding it somewhere is one) which requires that ψ be quadratically integrable, and (2) ψ must be continuous and single-valued.

The operator in equation 2.9 that returns the energy of the system, E, is the eigenvalue called the Hamiltonian operator, H. Thus, we can write

$$H\psi = E\psi \tag{2.10}$$

which is known as the *Schrödinger equation*. This Hamiltonian operator, H, is composed of five terms: (1) the kinetic energy of the nuclei, (2) the kinetic energy of the electrons, (3) the potential energy of nuclear repulsion, (4) the potential energy of electronic repulsion, and (5) the potential energy of nuclear-electronic attraction, which are shown in equation 2.11.

$$H = -\sum_i \frac{\mathrm{h}^2}{2m_e}\nabla_i^2 - \sum_k \frac{\mathrm{h}^2}{2m_k}\nabla_k^2 - \sum_1 \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \qquad (2.11)$$

where i and j run over all electrons, k and l run over all nuclei, $\hbar$ is Planck's constant divided by 2π, m_e is the mass of the electron, m_k is the mass of the nucleus k, ∇^2 is the Laplacian operator, e is the charge of the electron, Z is an atomic number and r_{ab} is the distance between particles a and b. Thus, the Hamiltonian operator contains pair-wise attraction and repulsion terms, implying that no particle is moving independently of all of the others (the term "correlation" is used to describe this interdependency). In order to simplify this complex system a number of approximations are made, which are discussed in the following sections.

2.3.1.1 The Born-Oppenheimer Approximation

The Born-Oppenheimer Approximation (BOA) suggests that because the nuclei are much heavier than the electrons, the nuclei may be treated as stationary in the field of moving electrons. Thus, the kinetic energy of the nuclei (term 1, equation 2.11) becomes zero and the potential energy of nuclear repulsion (term 3, equation 2.11) can be replaced by a constant which is dependent only upon the fixed positions of the nuclei.

Therefore, three terms are left in equation 2.11, and just one of them presents any real difficulty for modelling purposes: the potential energy of electronic repulsion (term 4, equation 2.11). If electrons did not repel one another then each electron would behave as an individual and the equation could be solved as a series of N one-electron equations, the summation of which would give the overall solution.

However, the repulsion energy between two electrons, which is around 323 kcal mol^{-1} (1350 kJ mol^{-1}) at a separation of 1 Å, is too high to be ignored so, instead of neglecting their repulsion, their correlation can be ignored. This leads to the most basic level of *ab initio* calculation, the Hartree-Fock (HF) or Self-Consistent Field (SCF) theory.

2.3.1.2 Hartree-Fock Theory

Hartree-Fock (HF) theory is fundamental to much of electronic structure theory. It is the basis of molecular orbital (MO) theory, which postulates that the motion of each electron can be described by a single-particle function (orbital) which does not depend explicitly on the instantaneous motions of the other electrons. This theory was developed to solve the electronic Schrödinger equation that results from the time-independent Schrödinger equation (equation 2.11) after invoking the BOA. In HF theory all the electrons are moving in a static potential created by the other electrons and nuclei, and are behaving like individual non-interacting electrons, feeling the repulsion of each other only in an averaged sense. Solutions to the N one-electron Schrödinger equations are found in an iterative fashion, and for this reason this model is also often referred to as self-consistent field (SCF) theory. The theory offered, for the first time, an accessible way of solving the complex equations that govern the quantum state of matter, but as it ignores the correlation between electrons, the accuracy of the final results can be seriously affected.

This approximation (that the electrons do not interact with each other) means that it is possible to separate the Hamiltonian into one-electron components. Thus the total electronic wave function $\psi(r_1,r_2, \ldots, r_n)$, describing the motions of N electrons is just the product of N hydrogen atom wave functions (orbitals), $\psi_H(r_1)\ \psi_H(r_2)\ldots\ \psi_H(r_n)$. In general:

$$\psi_{HP}(r_1, r_2, ..., r_N) = \phi_1(r_1)\phi_2(r_2)...\phi_N(r_N), \tag{2.12}$$

which is known as the *Hartree Product*.

This functional form, although fairly convenient, fails to satisfy the *anti-symmetry principle* (exchange symmetry), which states that a wave function describing fermions should be anti-symmetric with respect to the interchange of any set of space-spin coordinates (fermions have three spatial coordinates, xyz, and an intrinsic spin coordinate: spin up ↑, and spin down ↓). This requirement of exchange symmetry means that for two electrons the simple product wave function shown in

equation 2.12 is not correct. If both electrons exhibited the same spin "up", the two-electron wave function must be:

$$\psi(r_1, r_2) = \psi_{a\uparrow}(r_1)\psi_{b\uparrow}(r_2) - \psi_{a\uparrow}(r_2)\psi_{b\uparrow}(r_1) \tag{2.13}$$

Thus if r_1 is set to r_2 (*i.e.* the two electrons are located in the same place) the wavefunction vanishes, thereby satisfying *Pauli's exclusion principle* which states that no two electrons can be assigned the same set of principle quantum numbers.

2.3.1.3 Møller-Plesset (MP) Perturbation Theory

We saw in the previous section that HF theory allowed the complex (inaccessible) set of quantum mechanical equations describing the electronic behaviour of matter to be solved as a series of (accessible) one-electron Hamiltonians. The one drawback with this approach, however, was the necessary neglect of electron correlation energy. The theory can be improved, however, and an approximation to the missing energy obtained, by applying a small "perturbation" to the HF function. Thus, in equation 2.10 the operator H would be $H = H_0 + H'$, where H_0 is the HF function and H' is the applied perturbation.

In conceptual terms this means adding virtual excitations to the electronic ground-state wavefunction. When the perturbation is carried out to the second order (MP2), the resulting calculation encapsulates about 90% of the electron correlation energy. More of the 'missing energy' can be obtained by invoking higher perturbations (e.g. MP3, MP4 etc) but at great computational cost. For this reason, other computational models have been sought. The most important alternative, density functional theory, is discussed in the following section.

2.3.1.4 Density Functional Theory

2.3.1.4.1 Hohenberg-Kohn theorems and Kohn-Sham formulations

The basic theorems of the density functional formalism were derived by Hohenberg & Kohn (1964), who introduced the *Hohenberg-Kohn (HK) theorem*. This theorem states that given a ground-state density $n_0(r)$ it is possible, in principle, to calculate the corresponding ground-state wavefunction $\psi_0(r_1,r_2\ldots,r_N)$. This means that ψ_0 is a functional of n_0 and consequently, all ground-state observables are functionals of n_0 too. If ψ_0 can be calculated from n_0 and vice versa, both functions are equivalent and contain exactly the same information. The crucial fact that makes this possible is that knowledge of $n_0(r)$ implies implicit knowledge of much more than that of an arbitrary function $n(r)$. Knowledge that a given function is a ground-state density implies knowledge of an extremely detailed subsidiary condition: this function represents the spatial distribution of probability of the lowest energy solution to an N-particle second-order differential equation, namely the Schrödinger equation (Capelle, 2003).

In summary, the authors essentially state and prove two theorems: (1) the electron density determines the external potential, and (2) the energy for the electron density assumes its minimum value for the correct electron density, if the admissible functions satisfy the condition that the integral of the electron density is equal to the number of electrons.

As it was explained in previous sections, the difficulty in solving the Schrödinger equation for a polyelectronic system derives from the electron-electron interaction term in the correct Hamiltonian. Like Hartree, Kohn and Sham (1965) also appreciated that things would be considerably simpler if the Hamiltonian operator was formalised for a non-interacting system of electrons. Such a Hamiltonian can be expressed as a sum of one-electron operators, and has eigenvalues that are simply the sum of the one-electron eigenvalues. The most important part of the theory, however, is to take as a starting point a fictitious system of non-interacting electrons that have

for their overall ground-state density the same density as some real system of interest where the electrons do interact.

The Kohn-Sham equations look like standard HF equations, except that the exchange term is replaced with an exchange-correlation potential, V_{xc}, whose form is unknown. The exchange-correlation term is often split into a sum of two parts: one for the exchange effects and another for correlation effects. All that is needed, in principle, in order to solve the KS equations is a simple modification to standard HF computer codes, where the exchange contribution is replaced by the KS contribution. Like the methods based on Hartree-Fock theory described in the previous section, there exist a hierarchy of functionals of varying accuracy, which are described in the following two sections.

2.3.1.4.2 Local Density Approximation

The local density approximation (LDA) calculates the value of the exchange-correlation potential at some position *r* from the value of the density ρ at that position (i.e., the "local" value of ρ). In principle then, the only requirement on ρ is that it must be single-valued at every position, and it can otherwise be wildly ill-behaved. In practice, the only working functionals conforming to this definition are those that derive from analysis of the uniform electron gas, where the density has the same value at every position.

One of the most important deficiencies of the LDA exchange is that it does not have the correct asymptotic behaviour. Consequently, LDA found most success for systems where the electron density distribution fluctuates to only small extents (e.g. metallic systems). Improvements to these basic functionals gave rise to the Generalised Gradient Approximation (GGA) functionals, which have found considerable success for systems where the electron density fluctuations are much more pronounced, e.g. molecular systems.

2.3.1.4.3 Generalised Gradient Approximation

In a molecular system, the electron density is typically rather far from spatially uniform, so there is good reason to believe that the LDA approach will have some limitations. One obvious way to improve the correlation functional is to make it depend not only on the local value of the density, but on the extent to which the density is locally changing, i.e., the gradient of the density. Such an approach was initially referred to as 'non-local' DFT because the Taylor-expansion-like formalism implies reliance on values of the density at more than a single position. Mathematically speaking, however, the first derivative of a function at a single position is a local property, so the more common term in modern nomenclature for functionals that depend on both the density and the gradient of the density is 'gradient-corrected' or sometimes the 'generalised gradient approximation' (GGA).

The most popular GGA exchange functional to date has been one developed by Becke (1988). Usually abbreviated simply 'B', this functional has the correct asymptotic behaviour at long range for the energy density, and it incorporates a single empirical parameter, the value of which was optimised by fitting to the exactly known exchange energies of the six noble gas atoms He through Rn. There are many other alternative functionals that can be used (e.g. PW, Burke, Perdew and Wang, 1998).

A very popular GGA correlation functional, LYP, does not correct the LDA expression but computes the full correlation energy (Lee, Yang, and Parr 1988). It provides an exact cancellation of the self-interaction error in one-electron systems.

Typically in the literature, a complete specification of the exchange and correlation functionals is accomplished by concatenating the two acronyms in that order. Thus, for example, a BLYP calculation combines Becke's GGA exchange with the GGA correlation functional from Lee, Yang and Parr.

2.3.1.5 Basis sets

2.3.1.5.1 Localised basis sets

A basis set is a set of mathematical functions that are combined to approximate the wavefunctions for electrons in atoms and molecules. In other words, these basis functions are used in building up the description of molecular orbitals (an 'orbital' is a one-electron function). In *ab initio* molecular orbital (MO) theory, the basis functions can take many forms and they are not necessarily related to the atomic orbitals. Basis sets are calculated using two types of equations: *Slater type orbitals* (STOs) or *Gaussian type orbitals* (GTOs):

$$STO = \left(\frac{\zeta^3}{\pi}\right)^{0.5} e^{(-\zeta \cdot r)} \tag{2.14}$$

$$GTO = \left(\frac{2\alpha}{\pi}\right)^{0.75} e^{\left(-\alpha \cdot r^2\right)} \tag{2.15}$$

The STOs are not particularly amenable to implementation in molecular orbital calculations due to the difficulty in evaluating some of the integrals, particularly when the atomic orbitals are centred on different nuclei. ζ (zeta) controls the width of the orbital, so a large ζ gives a diffuse function. The basis sets derived from such orbitals are described as being localised as they are focused on atomic centres.

It is common in *ab initio* calculations to replace STOs by functions based upon Gaussians (i.e., GTOs). In the equation for a GTO (equation 2.15), α determines the radial extent of the Gaussian functions so that if α exhibits a small value, the function does not spread very far.

There are different classifications of basis sets: *minimal*, which is a representation that contains just the number of functions that are required to accommodate all the filled orbitals in each atom; *double-zeta*, which uses two basis

functions for each atomic orbital; *triple zeta*, which uses three basis functions for each atomic orbital, and so on for *quadruple-zeta* (QZ), *5Z*, *6Z*, etc. A very common type of basis set used is the so-called split-valence basis set, which uses one block of basis functions for each core atomic orbital, and a larger basis for the valence atomic orbitals [e.g., single-zeta is represented by 3-21G, which uses one block of three Gaussian functions to describe the core orbitals and another two blocks of Gaussian functions (denoted by the '2' and '1') to describe the valence orbitals]. The rationale for this approach is that the core orbitals, unlike the valence orbitals, do not affect chemical properties very much and vary only slightly from one molecule to another. Thus any computationally demanding improvements on the basis HF formalism (e.g. MP2 or DFT) can focus on the separate valence wavefunction only.

The minimal basis sets are well known to have several deficiencies. Some of these are due to the presence of electronegative atoms, such as oxygen or fluorine, in the molecule, which can have wavefunctions significantly different in a hetero-atomic environment compared to the homo-atomic state, from which the basis set is constructed. In addition, a minimal basis set cannot describe non-spherical aspects of the electronic distribution not attributed to the electronic ground state, such as a 'p-orbital' distribution for an 's'-block element.

These problems can be overcome by supplementing the basic Gaussian-derived basis set with two additional types of function: diffuse (denoted by a '+' in basis set nomenclature) and polarisation (denoted by a '*'), respectively. Diffuse functions are essential for the accurate description of species such as anions and molecules containing lone pairs and electronegative elements which have a significant amount of electron density away from the nuclear centres. This failure arises because the amplitudes of the Gaussian basis functions are rather low far from the nuclei. Polarisation functions are introduced into the basis set to essentially 'plug the gaps' in the electronic distribution of the ground atomic state necessary to achieve an accurate model to account for the formation of molecular bonds. A polarisation function is described by a higher angular quantum number than the ground state, and so correspond to 'p' orbitals for the 's' block elements (e.g. bonding in H_2) and 'd' orbitals for the 'p'-block elements (e.g. bonding in CH_4).

One of the main disadvantages of using localised basis sets, whether they are Slater-type, Gaussian or linear combination of Gaussian functions, is that they cause a mathematical inconsistency, which is ignored in the calculation of polyatomic systems. In order to explain the reason for the inconsistency, we can consider a chemical system AB, where A and B are two interacting fragments. We can express their interaction energy simply as the difference in energy between the complex AB and the individual energies of its fragments A and B. However, the description of the fragment A within the system AB is often improved by the basis functions of fragment B and vice versa, whereas it does not happen in the calculation of the isolated fragments A and B. This results in a "false" improvement of the general description of the AB system with respect to the description of A and B in isolation (Salvador-Sedano, 2001), giving rise to increased interaction energies. This effect, firstly described by Jansen and Ros (1969), is known as Basis Set Superposition Error (BSSE) (Liu *et al.*, 1973). Although there are many methods to minimise or take out the BSSE, such as the use of exact wave functions or the use of a set of functions centred at different points in the space to calculate the energies of the complex and the individual fragments, the most common method to eliminate the BSSE is known as the Counterpoise (CP) Correction, which was proposed by Boys and Bernardi (1970). The correction is carried out by calculating the interaction energy of a given AB system so the separate energies of the constituent fragments, A and B, are determined using the full set of basis functions used to calculate the energy of AB. In other words, for each fragment calculation, the electrons belonging to the other fragment were omitted and their nuclear charges set to zero (so-called 'ghost' atoms). Obviously, the BSSE correction to the calculated energies (e.g. binding energies) depends on the quality of basis set used to perform the calculation. At the MP2/cc-pVTZ theory level, the BSSE correction is roughly 30-50% of the total energy, whereas at MP2/cc-pVQZ theory level, the BSSE correction is around 10% (Rappé et al., 2000).

2.3.1.5.2 Delocalised basis sets

There exists an alternative to the localised basis set approach, where rather than confining electrons to exist at (or close to) atomic centres, electrons are treated as

(nearly) free particles that can exist anywhere within a given confined space, defined by a set of periodic boundary conditions (PBC's). Basis sets of this type are called delocalised, or plane-wave, and are often considered the most obvious choice to carry out calculations on periodic systems. The introduction of periodic boundary conditions (PBCs) has been extensively applied in the theoretical modelling of crystalline solids (Makov et al., 1995) because: (1) PBCs are a simple way to impose the boundary conditions in calculations of condensed matter. (2) PBCs are compatible with plane-wave expansions, which in turn allow for relatively simple calculations of forces in molecular-dynamics simulations, and (3) unified numerical schemes can be set up to consider both periodic and aperiodic systems.

A plane-wave basis set has many advantages over localised basis sets (Segall *et al.*, 2002): (1) it is continuous, (2) plane waves can always be added to improve the basis set, (3) there is a single convergence criterion, (4) plane waves are mathematically simple, and their derivatives are products in k-space, (5) plane waves do not depend on atomic positions and finally, (6) owing to the continuity of the plane wave basis set, there is no need to apply a correction for the BSSE (Davidson and Feller, 1986). Despite these advantages, there are two important disadvantages that go with the use of the plane-wave basis set: (1) the number of plane waves needed is determined by the greatest curvature of the wavefunction, and (2) as a consequence of the close distance between the wavefunctions of the valence electrons and the nucleus, many rapid oscillations are needed to maintain the orthogonality with the wavefunctions of the core electrons. These oscillations give rise to large kinetic energies and consequently, the wave functions are very expensive to model with plane waves.

This disadvantage can be compensated for by removing the electrons whose wavefunctions are localised in the core region and replacing it with an effective potential. This pseudopotential is constructed such that the wavefunctions outside the core region are unchanged. This rather drastic sounding approach dramatically reduces the number of plane waves required to represent the wavefunctions with no appreciable loss in computational accuracy for standard calculations, (Gillan, 1997)

which bears testament to the chemist's belief that core electrons do not contribute to the structure and bonding nature of a material.

2.3.1.6 Brillouin Zone

The difference between isolated molecules (i.e., gas phase) and periodic systems (i.e., solid state) is that isolated molecules have isolated energy levels, such as electronic, vibrational and rotational levels, whereas periodic systems have a continuum of energy states, such as electronic band structures and density of states. Therefore, the modelling of the solid state and its periodicity is more challenging than the modelling of aperiodic systems due to difficulties arising when describing the continnum of energy states.

There is a very useful theorem used to simplify the study of periodic systems, known as Bloch's theorem. Bloch's theorem makes use of the periodicity of a crystal to reduce the infinite number of one-electron wavefunctions to be calculated to simply the number of electrons in the unit cell of the crystal (or half that number if the electronic orbitals are assumed to be doubly occupied –that is, spin degenerate). By using Bloch's theorem, the problem of the infinite number of electrons has now been replaced by the problem of expressing the wavefunction as a function of an infinite number of reciprocal space vectors within the first Brillouin zone (BZ) of the periodic cell, *k*. The BZ, which is a complete classification of the energy levels of a crystal, contains an infinite number of points, also known as *k-points*. The electronic states of a system are allowed exclusively at a set of k-points, which are given by the boundary conditions that apply to the bulk solid. The theorem states that any wavefunction of a periodic system must be the product of a cell-periodic part and a wavelike part, in order to preserve the translational symmetry of the density. Consequently, the electronic wavefunctions at each *k*-point can then be expressed in terms of a discrete plane wave basis set. Each of the plane waves of the basis set have a kinetic energy associated to them, so that the plane waves with a smaller kinetic energy will generally be more important than those with a very high kinetic energy. By introducing a plane wave energy cutoff a basis set is created with a finite size. This

energy cut-off will give rise to an error in the calculated total energy of the system but in principle this error can be estimated and minimised by increasing the size of the basis set. This can be done by allowing a larger energy cut-off; the cut-off that is used in practice depends on the size of the system under investigation and on its constituent elements.

2.3.2 Complementary techniques: theory and experiment working together

During the work performed to complete this thesis, high-pressure single crystal X-ray diffraction experiments were carried out to study the compressibility effects on the crystal structures of small organic and biological molecules. However, the crystal structures obtained in most of the cases at high pressure were only partial due to many restrictions which arise from the experiment (see Section 2.2.3.1), so quantum mechanical calculations were performed on these experimental models in order to complete them. An example of this is the difficulty in locating the H-atom positions at high pressure, owing to the bulk of the diamond anvil cell used to carry out the experiments. Quantum mechanical calculations can optimise the 'guessed' H-atom positions to return reasonable positions for these atoms. Thus, in this process both techniques benefit: the partial structure obtained by experiment is completed, without cause to turn to costly time-consuming and expensive neutron diffraction techniques, and the ability of the computational modelling to reproduce the experimental parameters (lattice vectors and heavy atom coordinates) can be fully assessed. But there is more to the process than just structure optimisation. One could consider that the experimental techniques can only give us a "picture" of what the crystal structure of a compound looks like; the computational work can also give us energies and atomic charges (without having to perform experimental work) to go beyond this picture.

For the new molecular crystal structures presented in this thesis, experimental structures are used as input for the calculations, so that the atom positions and the cell parameters can be optimised to find the lowest energy structure. Additional simulations then enabled the calculation of sublimation, lattice and proton transfer

energies as well as individual energies of the hydrogen bonds present in the structure. Once the optimised structure is obtained, a Mulliken population analysis (MPA) can be carried out on the structure to obtain the Mulliken atomic charges of the atoms (Mulliken, 1955). The MPA is a partitioning scheme based on the use of density and overlap matrices for allocating the electrons of a molecular compound in some fractional manner among its various parts (atoms, bonds, orbitals). Once the Mulliken charges are calculated, these can be used to calculate the electrostatic contribution to the different energies previously obtained. In our work this was done in two ways: an Ewald sum (Berthaut, 1952; Allen *et al.*, 1987) or a simple application of Coulomb's law. An Ewald sum is recognised as being the best technique to calculate electrostatic interactions in a periodic system. It has two different components: the real space part (a combined assembly of point ions and Gaussian charges) and the reciprocal space part (second set of Gaussian charges centred on the point ions, nullifying the effect of the first set of gaussians). Therefore, the Ewald sum replaces a potentially infinite sum in real space by two finite sums: one in the real space and one in the reciprocal space. Despite its advantages, the energy obtained through the Ewald sum is the overall electrostatic energy of a system and therefore, if one were interested only in the electrostatic energy arising from the interaction between two particular atoms in the crystal structure, a different approach would be necessary. In this thesis, we therefore developed an in-house program called V_Coulombic, to calculate the Coulombic electrostatic contribution of an intermolecular interaction to the overall crystal energy. The program is based on the Coulombic potential (equation 2.16), which is used to describe the interaction between two point charges.

$$V = \left(\frac{1}{4\pi\varepsilon_0} \right) \frac{q \cdot q'}{r} \tag{2.16}$$

where V is the Coulombic potential in energy units, ε_0 is the electric permittivity of space, q and q' are two electric charges in coulombs, and r is the distance between the two charges. Therefore, taking the quantum mechanically optimised structure and set of Mulliken charges as input, we can obtain a full listing of the atom-atom electrostatic interaction energies.

In summary, the experimental and computational techniques used in this work fully complement each other. Through their combination we obtain a complete structure solution and understanding of the intermolecular interaction energies for molecular crystal structures at high pressure.

2.4 References

Allan, D. R. & Clark, S. J. (1999). *J. Phys. Rev. B*, **60**, 6328-6334.

Allen, M. P. & Tildesley, D. J. (1987). *Computer simulation of liquids*. Clarendon Press, Oxford, UK.

Angel, R. J. (2004). *J. Appl. Cryst.* **37**, 486-492.

Becke, A. D. (1988). *Phys. Rev.* **A38**, 3098.

Belonoshko, A. B.; Li, S.; Ahuja, R.; Johansson, B. (2004a). *J. Phys. Chem. Sol.* **65**, 1565-1571.

Belonoshko, A. B.; Simak, S. I.; Kochetov, A. E.; Johansson, B.; Burakovsky, L.; Preston, D. L. (2004b). *Phys. Rev. Lett.* **92**, 195701/1-195701/4.

Betteridge, P. W.; Carruthers, J. R.; Cooper, R. I.; Prout, K.; Watkin, D. J. (2003). *J. Appl. Cryst.* **36**, 1487.

Bernasconi, M.; Chiarotti, Guido L.; Tosatti, E. (1995). *Phys. Rev. B*, **52**, 9988-98.

Berthaut, F. (1952). *J. Phys. Rad.* **13**, 499-505.

Blessing, R. H. (1995). *Acta Cryst.* **A51**, 33-8.

Boese, R. & Nussbaumer, M. (1994). *International Union of Crystallography, Crystallographic Symposia,* **7**, 20-37.

Boyd, D. B. & Lipkowitz, K. B. (1982). *J. Chem. Ed.* **59**, 269-277.

Boys, S. F. & Bernardi, F. (1970). *Mol. Phys.* **19**, 553.

Bragg, W. H. & Bragg, W. L. (1915). *Trans. Roy. Soc. London*, **215**, 253-274..

Bragg, W. L. (1953). *Inst. Intern. Chim. Solvay, Conseil chim., 9th Conseil, Brussels*, 100-9.

Bruker-AXS (1999). *GEMINI*. Version 1.01. Bruker-AXS, Madison, Winsconsin, USA.

Bruker-AXS (2002). *SAINT*. Version 6. Bruker-AXS, Madison, Winsconsin, USA.

Bruker-AXS (1997-2001). *SMART*. Version 5.049-5.059. Bruker-AXS, Madison, Winsconsin, USA.

Burke, K.; Perdew, J. P.; Wang, Y. (1998). *Electronic Density Functional Theory. Recent Progress and New Directions*, Plenum Press, New York.

Burnham, C. J. & Xantheas, S. S. (2002). *J. Chem. Phys.* **16**, 1479-1492.

Capelle, K. (2003). *VIII'th Summer School on Electronic Structure of the Brazilian Physical Society*.

Chisholm, J. A.; Motherwell, S.; Tulip, P. R.; Parsons, S.; Clark, S. J. (2005). *Crystal Growth & Design*, **5**, 1437-1442.

Cramer, C. J. (2003). *Essentials of Computational Chemistry: Theories and Models*, Wiley Ed., England.

Crowfoot, D.; Bunn, C. W.; Rogers-Low, B. W.; Turner-Jones, A. (1949). *Chemistry of Penicillin*, Princeton Univ. Press.

Davidson, E. R. & Feller, D. (1986). *Chem. Rev.* **86**, 681.

Dawson, A.; Allan, D. R.; Parsons, S.; Ruf, M. (2004). *J. Appl. Cryst.* **37**, 410-416.

Drew, H. R. III. (1981). *The high resolution structure of DNA by single-crystal x-ray methods.*

Fortes, A. D.; Brodholt, J. P.; Wood, I. G.; Vocadlo, L. (2003). *J. Chem. Phys.* **118**, 5987-5994.

Gillan, M. J. (1997). *Contemp. Phys.*, **38**, 115-130.

Hinchliffe, A. *Modelling Molecular Structures* (2001). Second edition. Wiley Ed., England.

Hodgkin, Dorothy C. (1955). *Bull. Soc. Franc. Min. Cryst.* **78**, 106-15.

Hohenberg, P. & Kohn, W. (1964). *Phys. Rev. B*, **136**, 864.

Jansen, H. B. & Ross, P. (1969). *Chem. Phys. Lett.* **3**, 140.

Klug, A.; Finch, J. R.; Franklin, R. R. (1957). *Nature,* **179**, 683-4.

Koepke, J.; Dieterich, W.; Glinnemann, J.; Schulz, H. (1985). *Rev. Scient. Instrum.* **56**, 2119-22.

Kohn, W. & Sham, L. J. (1965). *Phys. Rev.* **140**, A1133.

Leach, A. R. (2001). *Molecular Modelling: Principles and Applications*. Prentice Hall, England.

Lee, C.; Yang, W.; Parr, R. G. (1988). *Phys. Rev.* **B37**, 785.

Lipps, G.; Weinzierl, A. O.; von Scheven, G.; Buchen, C.; Cramer, P. (2004). *Nat. Struct. Mol. Biol.* **11**, 157-162.

Liu, B. & McLean, A. D. (1973). *J. Chem. Phys.* **59**, 4557.

Makov, G & Payne, M. C. (1995). *Phys. Rev.* **B51**, 4014-4022.

Mantz, Y. A.; Chen, B.; Martyna, G. J. (2005). *Chem. Phys. Lett.* **405**, 294-299.

Merrill, L. & Bassett, W. A. (1974). *Rev. Sci. Instrum.* **45**, 290-294.

Meyer, M. (2002). *Private Communication*.

Miletich, R.; Allan, D. R.; Kuhs, W. F. (2001). *Reviews in Mineralogy & Geochemistry (High-Temperature and High-Pressure Crystal Chemistry)*, **41**, 445-519.

Mulliken, R. S. (1955). *J. Chem. Phys*. **23**, 1833.

Rappé, A. K. & Bernstein, E. R. (2000). *J. Phys. Chem.* **A104**, 6117-6128.

Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. (1992). *Rev. Mod. Phys.* **64**, 1045-97.

Sadle, J. & Lapinski, L. (1986). *THEOCHEM*, **32**, 233-40.

Salvador Sedano, P. *PhD Thesis*. University of Girona (2001).

Sheldrick, G. M. (1998). *NATO ASI Series, Series C: Mathematical and Physical Sciences*, **507**, 401-411.

Schwarz, U.; Grzechnik, A.; Syassen, K.; Loa, I.; Hanfland, M. (1999). *Phys. Rev. Lett.* **83**, 4085-4088.

Segall, M. D.; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. (2002). *J. Phys. Condens. Matter*, **14**, 2717-2744.

Steere, R. L. & Schaffer, F. L. (1958). *Biochim. Biophys. Acta,* **28**, 241-6.

Sterer, E.; Pasternak, M. P.; Taylor, R. D. (1990). *Rev. Scient. Instrum.* **61**, 1117-19.

Sukihara, T. (1990). *Nippon Kessho Gakkaishi* **32**, 77-85.

Tinland, B. (1968). *Chem. Phys. Lett.* **2**, 433-444.

Yen, J. & Nicol, M. (1992). *J. Appl. Phys.* **72**, 5535-8.

Chapter 3

Structural and Computational Study of Cyclopropylamine

3.1 Introduction

The main features of the hydrogen bond, which determine its strength, are the donor-hydrogen to acceptor distance and the angle formed between them. The strength of the hydrogen bond increases as the donor H atom to acceptor distance decreases and as the angle donor-H atom-acceptor gets closer to 180°. The electronegative character of the donor and acceptor atoms plays an important role in the strength of hydrogen bonding. Thus, the more electronegative the atoms, the stronger the hydrogen bond. Typical distances for strong hydrogen bonds are between 1.2 and 1.5 Å, with angles between 170 and 180°. An example of such a bond is F-H…F. However, the type of hydrogen bond reported in this work is the N-H…N hydrogen bond, which is classified as moderate to weak (Desiraju *et al.*, 2001; Steiner, 2002), with distances between 1.5 and 2.2 Å and angles greater than 130°. This type of hydrogen bond is commonly present in amines and their derivatives, and it is found in cyclopropylamine. Pressure has a significant effect not only on the hydrogen bond but on other intermolecular interactions present in the crystal structure such as van der Waals forces and steric effects. Consequently, when pressure is applied to the structure, the molecules are displaced from their equilibrium positions to form a generally more compact packing. These changes are driven by the competition between the formation of hydrogen bonding, minimising steric effects and the necessity of achieving a maximum coordination index, and an optimum filling of space with an accompanying lowering of the density of the crystal structure. Hydrogen bond distances are, consequently, compressed (or expanded) as a consequence of pressure. Therefore, the study of how the properties of the hydrogen bonding are influenced by pressure might enhance our understanding of the hydrogen bond and, in turn, the crystal structures that these molecular systems adopt.

As well as inducing changes in the geometry of intermolecular interactions, pressure can also affect the crystal structures of small-molecule systems leading to the formation of previously unobserved polymorphs. We have already investigated the high-pressure polymorphism of a range of fundamental small-molecule systems including simple monoalcohols and carboxylic acids (for example, methanol, ethanol,

formic acid, acetic acid and propionic acid), ketones (such as acetone), and the oxoacids (such as sulfuric acid) (Allan *et al.*, 1998, 1999a-c, 2000, 2001, 2002). We have also investigated more complex molecular materials such as pharmaceutical compounds [e.g. paracetamol (Fabbiani *et al.*, 2003)].

We are now extending this work to include other monofunctional systems such as the amines and in the first part of this chapter we report a high pressure polymorph of cyclopropylamine, which has been determined at 1.2 GPa using single-crystal X-ray diffraction techniques. The new phase, Phase II, crystallises at high pressure in the orthorhombic space group *Pbca*, with one molecule in the asymmetric unit and eight molecules in the unit cell. The previously reported low-temperature Phase I structure crystallises in the rhombohedral space group *R3c*. A detailed analysis and comparison of the hydrogen bonding in both polymorphs reveals substantial differences. The molecules at high-pressure are no longer arranged in $R_3^3(6)$ rings but, instead, they form zigzag chains of molecules along the a-axis, expressed as *C(2)* in the graph set notation (Bernstein *et al.*, 1995).

The second part of the chapter is dedicated to the computational study of the two different polymorphs of cyclopropylamine at ambient and high pressure. The same computational method, previously used to investigate the crystal structure of ammonia and urea (Morrison *et al.*, 2003), is applied to the Phase I and Phase II crystal structures of cyclopropylamine with two goals: to further test the method and, to study the energetics of the two systems. The calculated energies for the N-H…N hydrogen bonds present in the Phase I and Phase II crystal structures of cyclopropylamine are expected to be of the order of < 4 kcal mol^{-1} (17 kJ mol^{-1}) (Steiner, 2002), due to the contribution to the crystal structure from dispersive forces, such as van der Waals interactions. One of the aims of this investigation is to be able to compare the resulting calculated energies for this uncharged system with the studies carried out in amino acids, which are zwitterionic systems, and which will be explained in the following chapters. Finally, the hydrogen bonding in the Phase I and Phase II crystal structures of cyclopropylamine will be investigated to complete the

"picture" of these systems obtained via experimental work, and get a better understanding of their stability and hydrogen bonding.

3.2 Experimental study of cyclopropylamine

3.2.1 Crystal growth

Liquid cyclopropylamine was loaded and pressurised in a Merrill-Bassett diamond anvil cell (DAC; Merrill *et al.*, 1974) that was equipped with 600 μm culet diamonds and a tungsten gasket. After the nucleation of several crystallites on pressure increase, the temperature was cycled close to the melting point, in order to reduce the number of crystallites until only one crystallite remained. Finally, a single crystal was obtained at 1.2 GPa, after the cell was allowed to cool to ambient temperature.

3.2.2 High pressure X-ray Crystallography

The DAC was then mounted and centred on a Bruker SMART APEX diffractometer (graphite monochromated Mo $K\alpha$ radiation) and a sequence of eight data-collection scans was initiated, following the data-collection strategy of Dawson *et al.* (2004). The *SMART* program (Bruker-AXS, 1997-2001, 1999) was used for data-collection control and, with a detector distance of 70 mm, 2θ was set at either +28° or -28° to provide maximum coverage. No beam stop was used as, with this selection of detector distance and 2θ, the primary beam does not impinge on the detector aperture. The eight scans were conducted as a sequence of 3σ frames that each had a range of 0.3° in ω and an exposure time of 30 s. The φ axis was fixed at either 90° or 270°, to ensure that the axis of the diamond-anvil cell was held parallel to the $\omega/2\theta$ plane so that absorption from the pressure-cell components was minimised and the maximum possible access of reciprocal space was achieved. The overall data-collection time was 15 h. The sample reflections were identified by hand with the aid of the *SMART* code, and an orientation matrix was determined using the *GEMINI* program (Bruker-AXS, 1999). Data integration and global-cell refinement

were performed with the program SAINT (Bruker-AXS, 2002). The program *ABSORB* (Angel, 2004) was used to apply a correction for the absorption caused by the DAC and gasket "shadowing", rejecting reflections for which either the incident or the diffracted beam was completely absorbed by the cell and resulted in shading of the detector. The remaining reflections were corrected for any residual absorption by the pressure-cell components with the program *SORTAV* (Blessing, 1995, 1997) and the transmission ranged from 1.017 and 1.273.

The structure was solved in *Pbca* using direct methods (*SHELXTL*; Sheldrick, 1997a) and subsequently refined against F^2. The crystal structure is composed of eight molecules in the unit cell with one in the asymmetric unit. The unit cell parameters are a = 5.0741(10), b = 19.7594(10) and c = 13.305(2) Å at 1.2 GPa. One of the most serious difficulties encountered in high-pressure crystallography is the limited volume of reciprocal space that can be sampled owing to shading by the body of the pressure cell. For the Merrill-Bassett cell used in this study, the volume of accessible reciprocal space is limited to approximately 40% of what would be expected for a sample collected on a fibre at ambient conditions to the same resolution. However, the completeness of the data set collected here is 77.9% to 2θ = 46.5°, despite the constraints of the pressure cell, owing to the relatively high symmetry of the P*bca* space group and the orientation of the crystal within the cell. All the hydrogen atoms were found in the difference map. A full anisotropic refinement with independent positional parameters for each atom was performed and a common isotropic displacement parameter was also modelled for the H atoms. The structure refined to R = 0.042%, R_w = 0.096% for 65 parameters and 344 data with F > 4σ(F). The final difference map extremes were 0.114 and -0.138 e Å^{-3} and the goodness of fit was 1.176. The refinement and experimental details are presented in Table 3.1.

Crystal data	
Chemical formula	C_3H_7N
Chemical formula weight	57.10
Cell setting, space group	Orthorhombic, *Pbca*
a, *b*, *c* (Å)	5.0741 (10), 9.7594 (10), 13.305 (2)
α, β, γ (°)	90.00, 90.00, 90.00
V (Å^3)	658.89 (19)
Z	8
D_x (Mg m^{-3})	1.151
Radiation type	Mo $K\alpha$
No. of reflections for cell parameters	844
θ range (°)	2.6-23.0
μ (mm^{-1})	0.07
Temperature (K)	293(2)
Crystal form, colour	Block, colourless
Crystal size (mm)	Not measured
Data collection	
Diffractometer	*CCD* area detector
Data collection method	ω scans
Absorption correction	Multiscan (based on symmetry-related measurements)
T_{min}	0.799
T_{max}	1.000
No. of measured, independent reflections	371, 371, 344
Criterion for observed reflections	I>2σ(*I*)
R_{int}	0.065
θmax (°)	23.3
Range of *h*, *k*, *l*	-4→ *h* →4 -10→ *k* →10 -13→ *l* →13
Refinement	
Refinement on	F^2
R[F^2>2σ(F^2)], *w*R(F^2), *S*	0.042, 0.096, 1.18
No. of reflections	371
No. of parameters	65
H-atom treatment	Mixture of independent and constrained refinement
Weighting scheme	$w = 1/[\sigma^2(F_0^2) + (0.0403\mathrm{P})^2]$, where $P = (F_0^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	<0.0001
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.11, -0.14

Computer programs used: SMART (Bruker-AXS, 1997-2001), *SHELXTL* (Sheldrick, 1997a), *SHELXS97* (Sheldrick, 1997b), *SHELXL97* (Sheldrick, 1997c).

Table 3.1. Experimental details for the data collection of cyclopropylamine at 1.2 GPa.

3.2.3 Results and discussion

3.2.3.1 Analysis of the low temperature (Phase I) structure of cyclopropylamine

Although the low-temperature structure of cyclopropylamine has already been reported (de Boer *et al.*, 1986; CSD refcode FIGYID), a complete description has not been given of how the molecules are arranged in the crystal structure or the hydrogen bonding. It is important, therefore, to describe the structure at low temperature in detail, in order to compare it to the crystal structure determined at high pressure.

Phase I of cyclopropylamine crystallises at 170.5 K in the rhombohedral space group *R3c*, with 18 molecules in the unit cell and one in the asymmetric unit. The hexagonal unit cell parameters are $a = 18.784(2)$ and $c = 5.494(2)$ Å[1]. The carbon atoms of the molecule lie in the same plane, parallel to {-1 7 –1} and the nitrogen is rotated out of this plane by 147.3º, with respect to the carbons. There are two different N-H…N hydrogen bond interactions (Table 3.2) present in the structure which are formed by each of the two hydrogen atoms in the amino group. The N-H6…N interaction forms three-membered rings, or trimers, stacked along the crystallographic c-axis, with alternate trimers rotated by 61.8° (Figures 3.1 and 3.2). This contact would be defined as $R_3^3(6)$ in the graph set notation, as the three-membered rings are apparently formed by three donors and three acceptors. However, the distance of this interaction is significantly longer, by ~ 0.2 Å, than the normal range for N…N hydrogen bonds and, although the N-H6…N bond angle (150°) would be within expected limits, it is unlikely to offer a significant contribution to the overall intermolecular bonding in the structure. The trimers are bridged by shorter N-H7…N hydrogen bonds, which form *C(2)* zigzag chains of molecules along the crystallographic *c* direction (Figure 3.3).

[1] All the structural data given in this chapter for the low temperature structure of cyclopropylamine were taken from the published structure by de Boer *et al.* (1986), which does not include e.s.d's on distance and angle values.

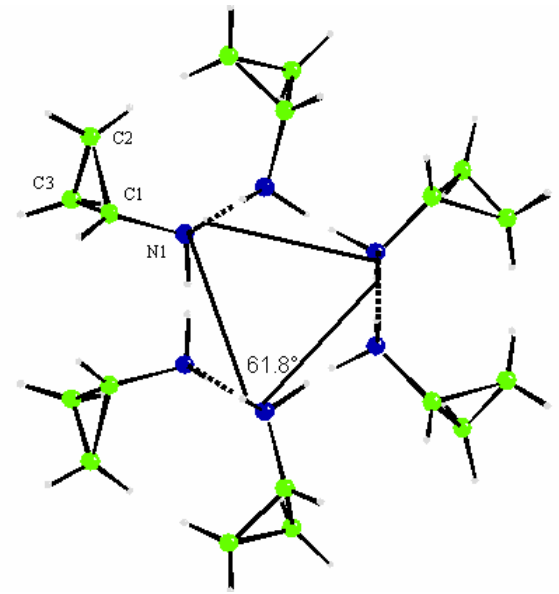


Figure 3.1: A view of the hydrogen bond pattern formed by two coupled $R_3^3(6)$ trimers in the low-temperature Phase I of cyclopropylamine (de Boer *et al.*, 1986). The trimers are formed by long N-H6…N interactions (shown by the solid lines) and alternate trimers are rotated by 61.8° with respect to their neighbours. The trimers are bridged by shorter N-H7…N hydrogen bonds.

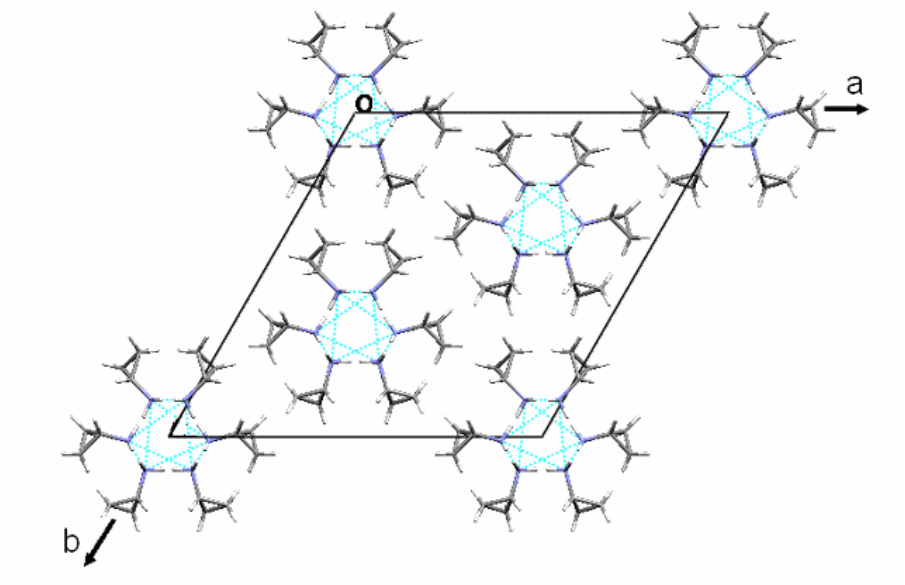


Figure 3.2: Projection down the crystallographic *c* axis of the low-temperature Phase I crystal structure of cyclopropylamine (de Boer *et al.*, 1986).

Interaction	D(N-H…N)/Å	D(N…N)/Å	Angle(N-H…N)/º	Angle(N…N-C)/º
Phase II				
N-H7…N^{i}	2.30(3)	3.166(3)	167(2)	117.5
Phase I				
N-H6…N^{ii}	2.711	3.541	150	130
N-H7…N^{iii}	2.285	3.230	177	111

Symmetry codes: (i) –1/2+x,y,1/2-z; (ii) –x+y,-x,z; (iii) x,x-y,-1/2+z.

Table 3.2: Selected hydrogen-bond distances (Å) and bond angles (º) for the high-pressure Phase II of cyclopropylamine from the experimental studies at 1.2 GPa and for the low-temperature Phase I of cyclopropylamine taken from the reported structure by de Boer *et al.* (1986).

An alternative way of describing the crystal structure is by examining an individual molecule and its neighbouring environment. By doing so, an octahedron is formed by molecules and their closest neighbours (Figure 3.4). Each polyhedron is formed by six molecules and the N atom of each of the molecules has a nearest N atom at a distance of 3.230 Å, a pair of N atoms at 3.541 Å and a single N atom at 3.629 Å. The octahedra are coupled to form infinite columns along the crystallographic c-direction where each of the octahedra share one face and follow an alternating sequence of parallel and perpendicular octahedra, with respect to their axial directions. Overall, each molecule is coordinated by 14 nearest neighbours.

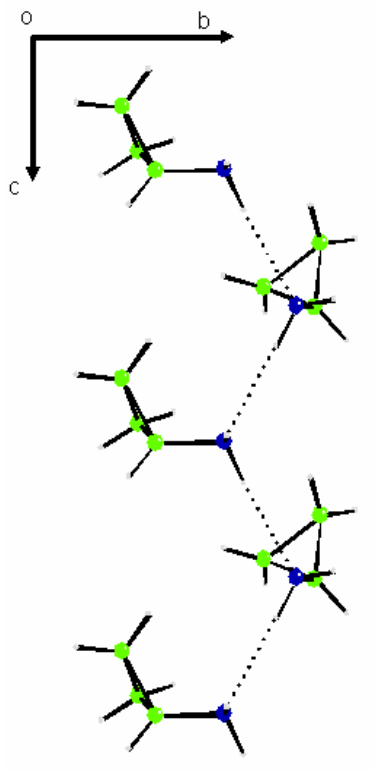


Figure 3.3: A view of the $C(2)$ chains, formed by the N-H7…N hydrogen bond, which run parallel to the *c*-axis direction in the low temperature Phase I crystal structure of cyclopropylamine (de Boer *et al.*, 1986).

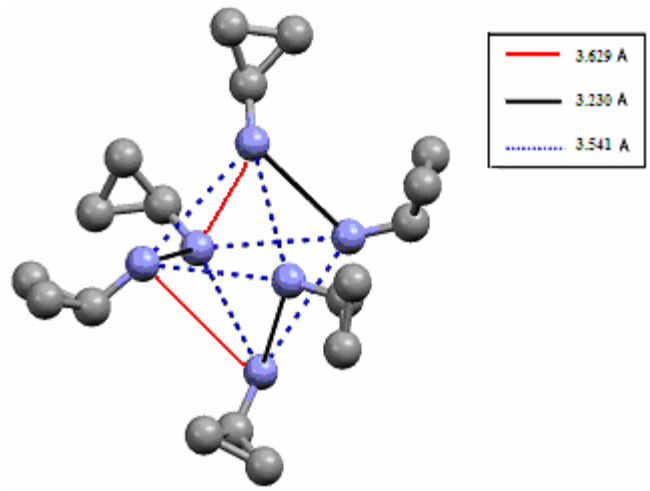


Figure 3.4: Octahedron formed by the arrangement of molecules linked by the N-H…N hydrogen bond and the weak N-H…N contact in the low temperature-phase of cyclopropylamine (de Boer *et al.*, 1986). Only the non-H atoms are shown. The bond distances (edges) of the polyhedron are displayed.

3.2.3.2 Analysis of the high-pressure (Phase II) structure of cyclopropylamine

At 1.2 GPa the N atom is rotated 148.7º out of the plane formed by the atoms C1, C2 and C3, which form the ring of C atoms. The cyclopropylamine Phase II structure at high pressure is simply formed by molecules connected in a zig-zag pattern, *via* a *cis*-N-H…N hydrogen bond, forming sets of infinite chains [N…N distance 3.166(3) Å], expressed as $C(2)$ in graph set notation (the repeating unit contains one hydrogen bond donor and one acceptor). The same molecular conformation found in Phase I is also found in the high pressure Phase II structure.

These chains lie parallel to the crystallographic *a*-axis, as can be seen from Figures 3.5 and 3.6. Only one of the two H atoms, H7, from the amino group participates in the hydrogen bonding. All the molecules of cyclopropylamine lie in the same side of the plane formed by the hydrogen bond network (Figure 3.5).

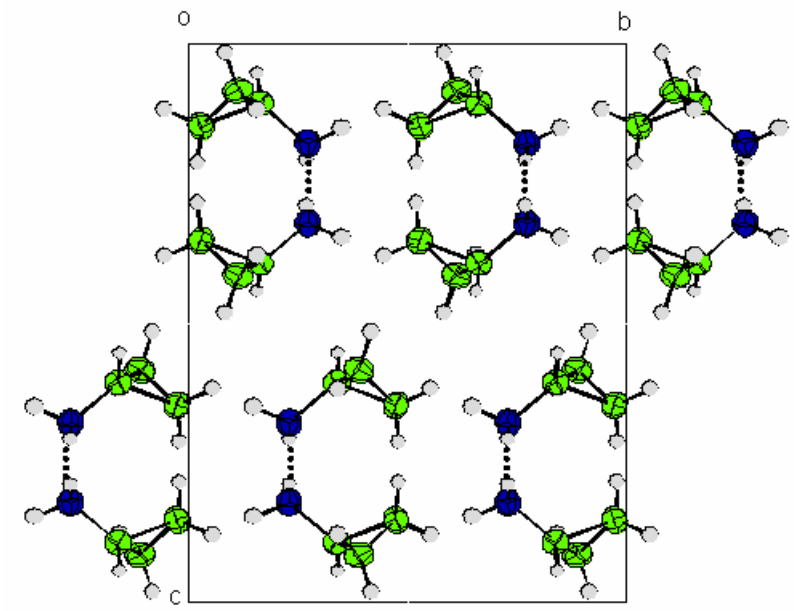


Figure 3.5: An *a*-axis projection of the high-pressure, P*bca*, structure of cyclopropylamine at 1.2 GPa. The crystallographic *b*-axis is horizontal while the *c*-axis is vertical. The molecules are linked via *cis* hydrogen bonds.

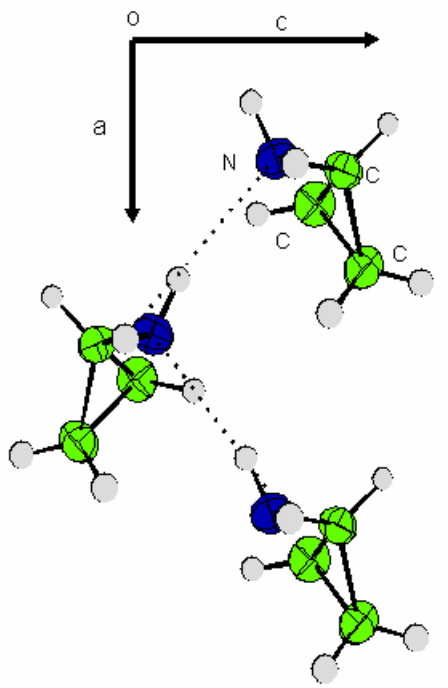


Figure 3.6: One of the crystallographically *C(2)* unique hydrogen-bonded chains in cyclopropylamine at 1.2 GPa viewed along the *b*-axis. The molecules adopt an alternating 1–1–1 sequence.

From the view looking down the hydrogen bond (N-H…N), the dihedral angle formed by the atoms involved in its formation (C-N-N-C) is 71º, which is relatively close to the 60º required for an ideal *gauche* conformation. Details about the distances and angles of the hydrogen bond can be found in Table 3.2.

The packing motifs adopted by monoalcohols, *R*OH, (Brock *et al.*, 1994) arise from a compromise between the packing requirements of the relatively bulky *R* groups present in the molecules and the necessity of the hydroxyl groups to get close enough to form the hydrogen bonding. If the *R* groups are relatively small, the molecules containing the hydroxyl groups will be related by translational symmetry (glide plane or 2_1 screw axis), forming an approximately coplanar alternating sequence about the central hydrogen-bonded core. For bulkier *R* groups the molecules often cannot adopt a simple arrangement owing to steric effects and instead they form

chains along three-, four- and sixfold screw axes. Finally, if the *R* groups occupy an even greater volume, cyclic dimer, trimer, tetramer or hexamer rings can be formed.

Thus, the structures of the two cyclopropylamine polymorphs can be compared to the structures of the monoalcohols at both low temperature and high pressure, which, in general, form molecular chains differing in the arrangement of the molecules about the chains. For example, the crystal structure of ethanol at ambient pressure (Jönsson, 1976) forms infinite chains where the molecules are connected by *cis* and *trans* O-H…O hydrogen bonds. However, at high pressure, the molecules of ethanol form linear hydrogen bonded chains, where the molecules are linked in each chain with their methyl groups aligned in the same direction along the *b*-axis (Allan *et al.*, 1999*a*, 2001). Other similar cases are the low temperature and high pressure polymorphs of methanol (Tauer & Lipscomb, 1952; Narten & Habenschuss, 1984; Torrie *et al.*, 1989; Allan *et al.*, 1998), cyclobutanol (McGregor *et al.*, 2005) and phenol and mono-fluorophenols, such as 2-chlorophenol and 4-fluorophenol (Oswald *et al.*, 2005). All the mono-fluorophenols as well as phenol and cyclopropylamine crystallise under pressure in low-symmetry space groups with the molecules disposed about 2_1 screw axes. Thus, cyclopropylamine presents a polymorphism similar to that found in alcohols, where pseudo-helices and ring motifs are often observed at low temperature. However, at high pressure they tend to form hydrogen bonded chains with a simple alternating sequence.

3.2.4 Comparison of the low temperature Phase I and high pressure Phase II crystal structures of cyclopropylamine

3.2.4.1 Hirshfeld Surfaces

The program *Crystal Exlorer* (Grimwood *et al.*, 2004) makes use of Hirshfeld surfaces (Hirshfeld, 1977) to partition crystal space in molecular crystals so that the packing modes and intermolecular interactions can be explored (McKinnon *et al.*, 2004). We have used this program to visualise the packing behaviour in Phase I and

Phase II of cyclopropylamine in order to make a more detailed comparison between them.

Hirshfeld surfaces (Hirshfeld, 1977) for the low-temperature and high-pressure polymorphs of cyclopropylamine are shown in Figure 3.7, with the H atoms of the amino group at the top pointing away from the viewer in each case. Only one of the two H atoms of the amino group in the high-pressure Phase II polymorph actively participates in the hydrogen bond. This is shown by the orange-red region on the d_e surface adjacent to the N atom (where d_e is the distance to the nearest atom centre exterior to the surface). This situation is also observed in the even members of the diamines (e.g. 1,2-ethanediamine, 1,4-butanediamine, 1,6-hexanediamine), as it was explained by McKinnon *et al.* (2004). Hirshfeld surfaces for the high-pressure polymorph show that there is clearly only a single interaction that could be classified as a hydrogen bond (interaction 1 in Figure 3.7). However, the Hirshfeld surfaces for the low temperature polymorph show a strong contact and a weak contact (interactions 1 and 2 in Figure 3.7) present in the structure, which is common to the odd members of the diamines (e.g. 1,3-propanediamine and 1,5-pentanediamine).

The two-dimensional fingerprint plots (plot of d_i *vs* d_e, where d_i is the distance to the nearest atom centre interior to the surface) for the two polymorphs stress the systematic differences between the two structures (Figure 3.8). One of the main differences is that the voids (upper region of the plots in Figure 3.8) are more compact in the high-pressure Phase II crystal structure than in the low temperature Phase I structure indicating that the packing is more efficient. The second difference is that both structures present relatively short hydrogen bonds (the distances labelled 1 in Figure 3.8), one of which is slightly shorter in the case of the Phase II polymorph. Finally, the Phase I structure has two significant hydrogen bonds, while the high-pressure structure has only the one, short, hydrogen bond with the remaining amino H atom not involved in bonding. It can be seen that the second, longer, hydrogen bond of the low temperature Phase I polymorph is weak in nature as it does not form part of the sharp spikes in the Hirshfeld fingerprint, which represent the strong hydrogen bonds (the distance labelled 2 in Figure 3.8).

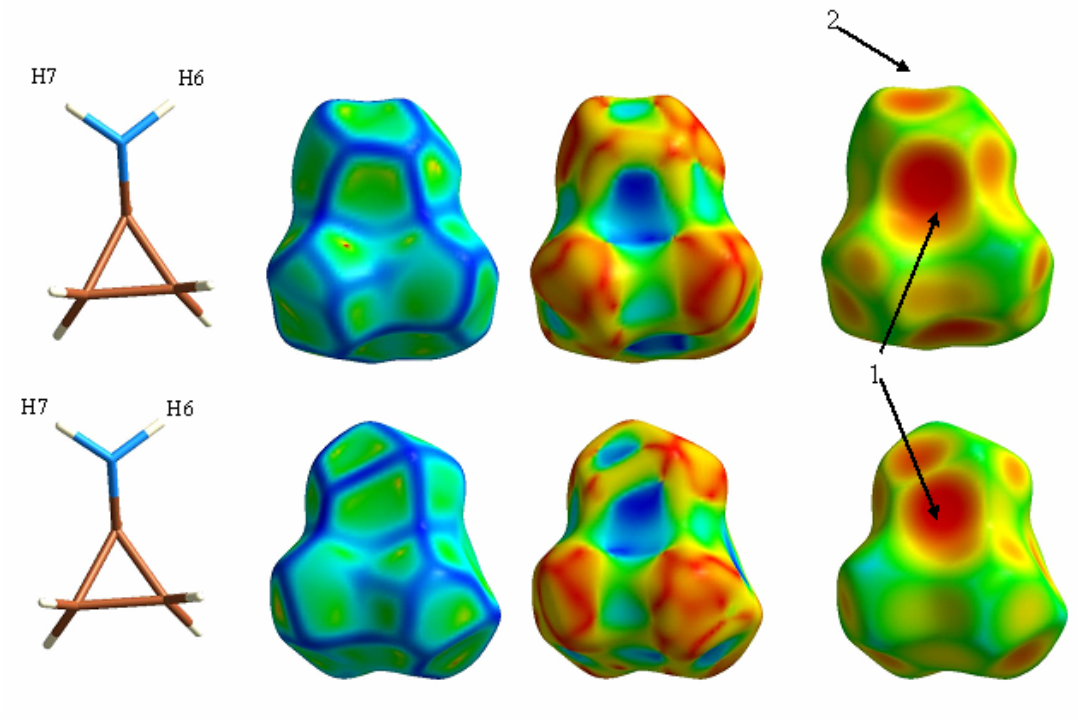


Figure 3.7: Hirshfeld surfaces for the low temperature Phase I (top) and high-pressure Phase II (bottom) polymorphs of cyclopropylamine. Each molecule is shown with the Hirshfeld surface mapped with curvedness (left), shape index (centre) and d_e (right), which is the distance to the nearest atom centre exterior to the surface [for this series mapped between 1.0 (red) and 2.5 Å (blue)]. The labels 1 and 2 correspond to the strong and weak hydrogen-bond interactions, respectively, present in the crystal structures of Phase I and Phase II. The NH_2 group of the cyclopropylamine molecule for both these projections is oriented in the same way for Phase I and Phase II with the N atom pointing towards the viewer and the H atoms pointing away from the viewer (as can be seen from the sketch of the molecule shown beside the Hirshfeld surfaces). The H7 and H6 atoms of the low temperature Phase I are involved in the formation of a short and a long hydrogen bond, respectively, whereas the H7 atom of the high pressure Phase II is the only H atom that participates in the formation of hydrogen bonding.

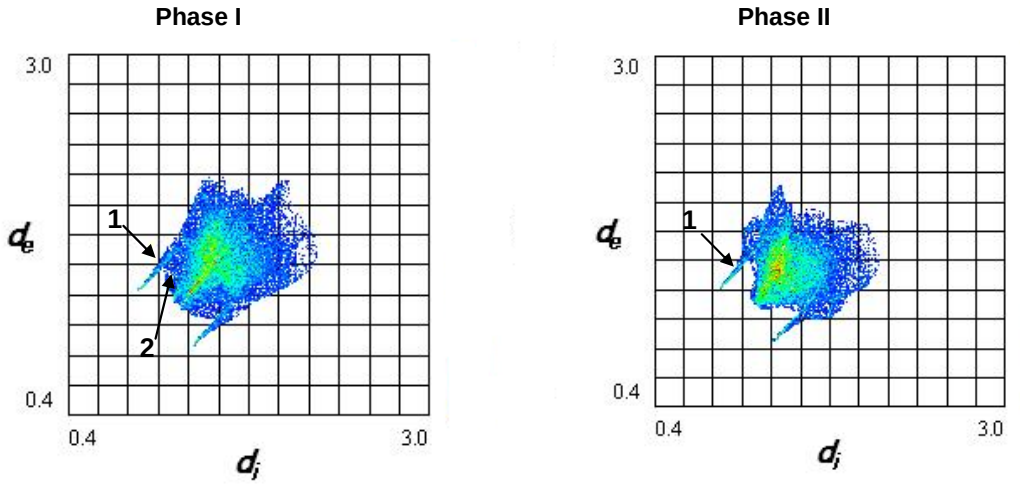


Figure 3.8: Two-dimensional fingerprint plots for the low temperature (left) and high-pressure (right) polymorphs of cyclopropylamine. The labels 1 and 2 correspond to the strong and weak hydrogen-bond interactions, respectively, present in the crystal structures of Phase I and Phase II. d_e and d_i are the distances to the nearest atom centre exterior and interior to the surface, respectively.

3.2.4.2 Topological analysis

Topological analyses of Phase I and Phase II can be useful to rationalise the more significant structural changes caused by pressure. Voronoi-Dirichlet polyhedra (VDPs) were used to analyse the molecular structure of the two polymorphs of cyclopropylamine (Blatov, 2004).

The calculation of molecular coordination numbers, and topological and geometrical analysis of the environment of the two different polymorphs of cyclopropylamine were carried out by using the *TOPOS4.0* program suite (Blatov *et al.*, 1999). Adjacent matrices were calculated using the program *AUTOCN* using the method of spherical sectors; the minimum solid angle of a Voronoi-Dirichlet polyhedron (VDP) face corresponding to an intermolecular contact was set to zero. Analyses of the VDPs were carried out with the program *ADS*, with the geometrical centres of the molecules (as opposed to their centres of gravity) as reference points. Coordination sequences were calculated out to three coordination spheres. The results

of the topological analysis of cyclopropylamine polymorphs are presented in Figure 3.9. The coordination sequences are 14–54–122 and 14–53–120 for the Phase I and Phase II crystal structures, respectively. Consequently, it can be seen that the structure of Phase II is closer to a perfect body-centred cubic structure (14–50–110) than that of Phase I. Nevertheless, both structures are significantly distorted from body-centred cubic.

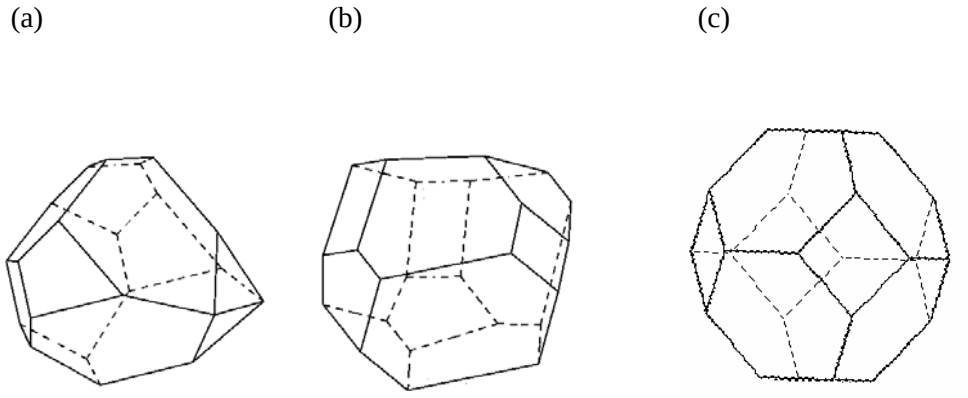


Figure 3.9: Lattice Voronoi-Dirichlet polyhedra in: (a) low temperature polymorph and (b) high-pressure polymorph of cyclopropylamine and (c) for perfect body-centred cubic packing. *Combinatorial types* (Peresypkina *& Blatov*, 2000) can be written $f/v-n$, where f is the number of faces and v the number of vertices; n is just an ordinal number to distinguish different VDPs where f and v are equal. These are: (a) 14/21-1, (b) 14/24-1 and (c) 14/24-1. The combinatorial types for (b) and (c) are therefore equivalent.

3.3 Computational study of cyclopropylamine

Following the experimental study of the solid state structures of cyclopropylamine, which was discussed in detail in the first part of this chapter, a computational study of the Phase I and Phase II crystal structures of cyclopropylamine was carried out. The aims of this computational investigation were:

a) to calculate the sublimation energy of both Phase I and Phase II structures, in order to understand the stability of the cyclopropylamine molecule in the two different structures, at ambient and high pressure,

b) to calculate the energy of the hydrogen bonds present in the two different solid state structures in order to compare their nature and properties,

c) to investigate the σ-bond cooperativity effect on the different N-H…N hydrogen bonds which hold the crystal structures at ambient and high pressure conditions.

3.3.1 Results and discussion: Phase I

3.3.1.1 Crystal structure of the low temperature (Phase I) polymorph of cyclopropylamine

The starting model to perform the quantum mechanical calculations for Phase I was taken from the crystal structure, published by de Boer *et al*. (1986). The rhombohedral structure of cyclopropylamine is formed by molecules linked *via* a short N1-H7…N1 hydrogen bond (2.285 Å, 177°), which forms zig zag molecular chains along the a-direction, and a long N1-H6…N1 intermolecular contact (2.711 Å, 150°), which forms trimers of molecules stacked parallel to the c-direction. Thus, both hydrogen atoms in the NH_2 group actively participate in the formation of hydrogen bonding (see Figures 3.1 and 3.3). Details of the geometry of the two intermolecular interactions can be found in Table 3.2.

3.3.1.2 Literature energies

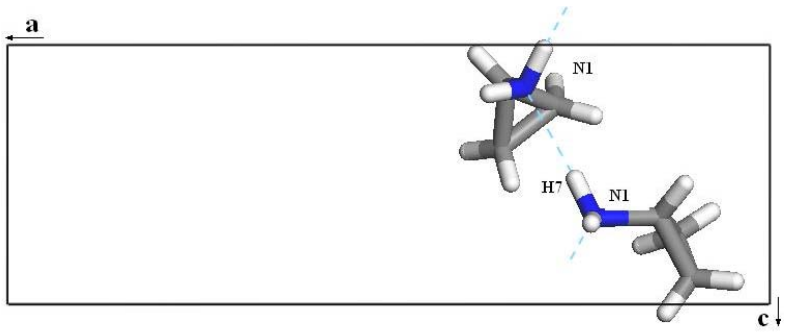
The literature cites no prior computational work having been performed on the crystal structure of cyclopropylamine. The results we obtained will therefore be compared with those obtained for Phase II, in order to benchmark the range of energies calculated for these systems.

3.3.1.3 Geometry optimisation

The crystal structure of the Phase I of cyclopropylamine contains one molecule in the asymmetric unit (Z'=1) and eighteen molecules in the unit cell (Z=18) [see Figure 3.2] (de Boer *et al.*, 1986). Each molecule is formed by eleven atoms, which means that the geometry optimisation would have to be performed on 198 atoms, and therefore would be very computationally expensive. Consequently, geometry optimisations for this polymorph were performed exclusively on models taken from the crystal structure in order to simulate the intermolecular interactions. The first model, named dimer 1, is formed by two molecules of cyclopropylamine connected *via* the short N1-H7…N1 hydrogen bond to create an infinite chain [Figure 3.10(a)]. Note periodicity can be 'switched off' for this interaction by stretching the model unit cell along the 'c' direction to 10 Å, offering the considerable attraction of investigating directly the effects of σ-bond cooperativity. The second model, trimer 1, is formed by three molecules forming an equilateral triangle (three-fold symmetry) connected by the long N1-H6…N1 contact [Figure 3.10(b)]. Note periodicity is not inherent in this interaction.

The results obtained by PW-DFT calculations via the CASTEP code (2002) for the geometry optimisations of the two models representing the hydrogen bonding present in the Phase I structure of cyclopropylamine can be found in Table 3.3, together with the experimental data for comparison. From this it can be seen that in general the computational results agree with the experimental input.

(a)



(b)

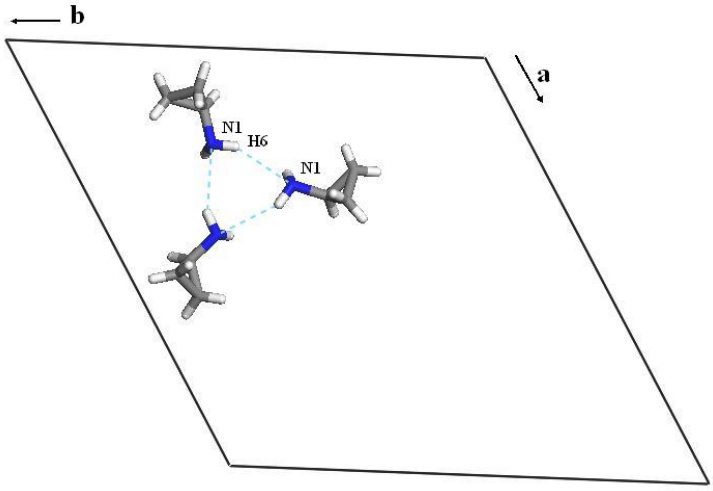


Figure 3.10: View of the two models chosen to perform the calculations on the Phase I structure of cyclopropylamine: (a) dimer 1 [*r*(H7…N1) = 2.285 Å] and, (b) trimer 1 [*r*(H6…N1) = 2.711 Å].

Parameters	**Experimental** [a]	**Calculated**	
	Solid	*Solid* (dimer 1, trimer 1)	*Supercell* (relaxed)
Lattice [Å, °]			
A	18.784	18.784	18.784
B	18.784	18.784	18.784
C	5.494	5.494	10
α=β,γ	90, 120	90, 120	90, 120
Z	18	2, 3	1
Volume [$Å^3$]	1678.8	1678.8	3528.4
Space/point group	*R3c*	*P1, P3*	*P1*
Geometry [Å, °]			
Dimer 1			
*r*N1…N1	3.230	3.236	—
*r*N1-H7	0.946	1.024	1.023
*r*N1…H_7	2.285	2.201	—
∠H7-N1-H6	110.3	108.9	108.6
∠H7-N1…N1	176.6	179.0	—
Trimer 1			
*r*N1…N1	3.541	3.466	—
*r*N1-H6	0.924	1.027	1.023
*r*N1…H6	2.711	2.555	—
∠H6-N1-H7	110.3	108.8	108.6
∠H6-N1…N1	150.1	147.7	—
Energy [kJ mol^{-1}]			
Sublimation	—	15.5, 6.4	21.9
Total energy [per molecule, eV]	—	-851.87005, -851.77575	-851.70951

[a] de Boer *et al.* (1986)

Table 3.3: Comparison between the experimental and calculated (PW-DFT) structures for the two intermolecular interactions present in the Phase I crystal structure of cyclopropylamine. There are two independent molecules in the dimer 1 calculations and therefore, the total energy reported here for a molecule is the average energy value for both molecules.

3.3.1.4 Calculated energies

All the calculated energies derived from the models mimicking the crystal structure of the Phase I polymorph can be found in Table 3.4. The calculated energies for dimer 1 were found to be 3.7 kcal mol^{-1} (15.5 kJ mol^{-1}) when the σ-bond cooperativity in the interaction is present (i.e. the hydrogen bond is modelled as periodic), and 2.7 kcal mol^{-1} (11.4 kJ mol^{-1}) when absent (i.e., the model periodicity is destroyed by increasing the 'c' unit cell vector to 10 Å). These results show that the σ-bond cooperativity effect has an important influence on the strength of the hydrogen bond, making the interaction 26% stronger. The increase in hydrogen bond strength due to periodicity obtained in this work found agreement with previously reported calculations on urea (ca. 25% enhancement) by Morrison *et al.*, 2003 and a value of 20% reported by Steiner (2002) for other molecular systems.

Model			CASTEP	
	r(D-H…A)	∠(D-H…A)	Energy/ eV	Energy/ kJ mol^{-1}
	r (D…A)	∠(C-D…A)		
Dimer 1	2.285	176.6		
Periodic			-851.87005	15.5
	3.230	110.8		
Aperiodic			-851.76880	11.4
Trimer 1	2.711	150.1		
Aperiodic	3.541	129.6	-851.77575	6.4
TOTAL				
Periodic				21.9
Monomer	—	—	-851.70951	—

Table 3.4: Energies obtained per molecule for the Phase I trimer, dimer and monomer models of cyclopropylamine, and the resulting energies per hydrogen bond (kJ mol^{-1}) (*r*/Å, ∠/°, D = donor, A = acceptor, C = Carbon atom).

The calculations performed on trimer 1 were more challenging than those performed on dimer 1 due to the existence of three-fold symmetry present in the trimer. When the geometry optimisation was performed in the absence of symmetry constraints the trimer motif was lost, and molecules realigned in an attempt to maximise dipole-dipole interactions. Nevertheless, even when the calculation was performed including the three-fold symmetry, the resulting optimised geometry was not quite as satisfactory as the optimised geometry obtained for dimer 1 (see Table 3.3). The calculated energy for trimer 1 was found to be 1.5 kcal mol^{-1} (6.4 kJ mol^{-1}) per interaction, which is half the strength of dimer 1. (Note there was no possibility of studying the influence of the σ-bond cooperativity on trimer 1 since the N-H…N interaction lacks periodicity in the crystal structure). We thus conclude that the Phase I packing arrangement displayed by cyclopropylamine is dominated by a strong periodic zig-zag hydrogen bonding interaction considerably enhanced by σ-bond co-operation, and a weaker non-periodic timer arrangement of approximately half the strength.

Since a geometry optimisation of the Phase I crystal structure would be too computationally expensive, an estimate of the sublimation energy of the system can be obtained instead by summing up the energies of the individual interactions present in the crystal structure. This estimate will of course only be reliable if the level of correlation between the two interactions is low. The number obtained, 5.2 kcal mol^{-1} (21.9 kJ mol^{-1}), is about the same order of magnitude found for ammonia [7(1) kcal mol^{-1} or 29(4) kJ mol^{-1}, Shipman *et al.*, 1976; 6.7 kcal mol^{-1} or 27.9 kJ mol^{-1}, Morrison *et al.*, 2003], which presents a similar N-H…N hydrogen bond network.

3.3.2 Results and discussion: Phase II

3.3.2.1 Crystal structure of the high pressure (Phase II) polymorph of cyclopropylamine

The starting model to perform the quantum mechanical calculations for Phase I was taken from the experimental crystal structure obtained at 1.2 GPa by Lozano-Casal *et al.* (2005). As described previously in the experimental discussion, the

orthorhombic structure is formed by molecules of cyclopropylamine linked via a short N1-H7…N1 hydrogen bond [2.30(3) Å, 167(2)°], which forms zig-zag molecular chains along the a-direction. Unusually only one of the hydrogen atoms in the NH_2 group actively participates in the formation of the hydrogen bonding network (see Figure 3.5). Details of the geometry of the intermolecular interaction can be found in Table 3.2.

The molecule of cyclopropylamine presents almost the same conformation in the Phase I and Phase II crystal structures (Figure 3.11). The angle formed by the N atom and the H atoms of the optimised NH_2 group is of 108.8 and 109.4° for Phase I and Phase II, respectively. The rotational angle formed between the optimised N atoms and the three-carbon ring is of 108.1 and 107.2° for Phases I and II, respectively. Thus, like the experimental study, the calculated structures also support the existence of one conformation of cyclopropylamine in both the ambient and high pressure crystal structures.

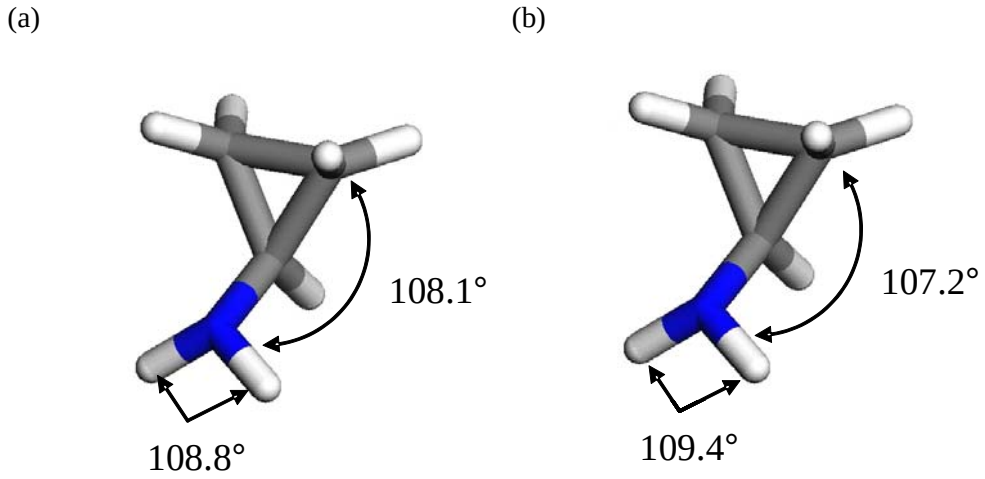


Figure 3.11: View of the conformation of the cyclopropylamine molecule in the crystal structures of (a) Phase I, and (b) Phase II, taken from the computational results. From the figures it is possible to see that the same conformation of the molecule is present in both polymorphs of cyclopropylamine.

A potential energy surface (PES) scan was carried out on a cyclopropylamine molecule, in order to understand why the NH_2 group of the molecule is involved only in the formation of one N-H…N hydrogen bond in Phase II (high pressure), whereas it forms two N-H…N interactions in Phase I (ambient pressure). The calculation was performed with the program Gaussian 98, allowing the cyclopropylamine molecule to optimise the atom positions, while the C1-C2-N1-H6 torsional angle was allowed to rotate in step-wise increments of 10° through 360°. Results show that the lowest energy conformation is compatible with the geometry present in the solid state. Moreover, a barrier height of around 22.6 kJ mol^{-1} needs to be overcome before another mimimum can be accessed (Figure 3.12). Thus we can conclude that cyclopropylamine is a small rigid molecule; twisting of the NH_2 group to participate in any further hydrogen bonding network is unlikely to occur. In terms of the molecular packing and as it was shown by the Voronoi-Dirichlet polyhedra [see section 3.2.3.2], the high pressure Phase II polymorph presents a crystal packing which is closer to a perfect BCC crystal structure than the low temperature Phase II polymorph. Considering the results obtained from the PES scan performed on the cyclopropylamine molecule, the difference in the packing can now be explained. In the high pressure structure only the stronger bond (i.e. more energetic) is retained to form a more compact molecular arrangement, giving up the weaker hydrogen bond which was present in the Phase II crystal structure forming trimers.

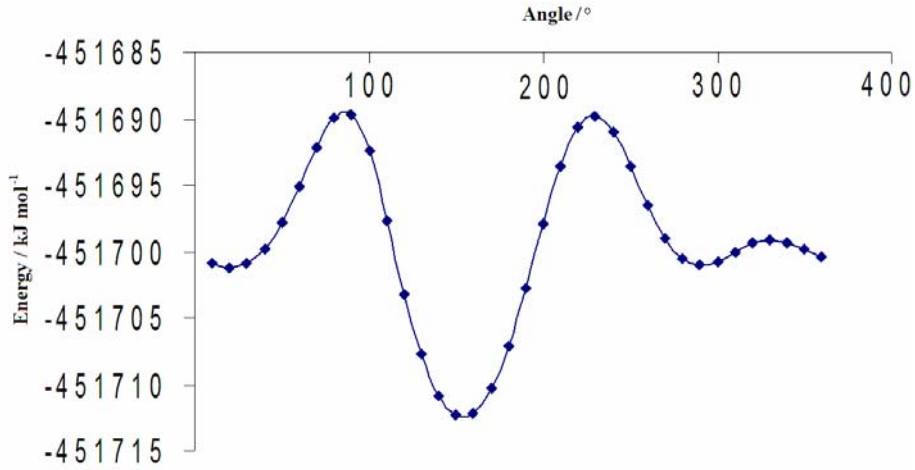


Figure 3.12: Graph showing the potential energy surface calculated for the C1-C2-N1-H6 torsion of the molecule of cyclopropylamine at 1.2 GPa.

3.3.2.2 Geometry optimisation

The results obtained by PW-DFT calculations for the geometry optimisation of the Phase II structure of cyclopropylamine can be found in Table 3.5, together with the experimental data for comparison. The results obtained are in close agreement with the experimental structure at 1.2 GPa. The difference in cell parameters are 1.2, 2.6 and 1.2% for 'a', 'b' and 'c' directions respectively. The cell volume increased, as expected for a GGA-DFT functional, by 5% compared to the experimental value. The simulation reported heavy and light atom positions in very close agreement with the input experimental fractional coordinates, with differences in the bonding parameters of the order of picometers [e.g. $r(\text{C1-N1})_{\text{exp}}$ = 1.434(3) Å, $r(\text{C1-N1})_{\text{calc}}$ = 1.428 Å].

In addition to the crystal structure geometry optimisation performed on Phase II, a geometry optimisation was also performed on a constructed model taken from the optimised geometry, named dimer 2, to mimic the N-H…N hydrogen bond present in the structure at 1.2 GPa (Figure 3.13), which is the only interaction that the PW-DFT calculation can model, in order to estimate other possible effects arising from interaction between neighbouring molecules in the crystal structure. The optimised geometry of the hydrogen bonded dimer is in close agreement with the experimental and computational geometries present in the crystal structure, and any differences in the hydrogen bond parameters are of the order of picometers (Table 3.5). Moreover, the difference in energy per molecule between this simplified model and the full crystal structure is only 0.01 eV, indicating that any other intermolecular interactions are clearly negligible. As the dimer 2 model is formed by two independent molecules (P1 symmetry) it is important to emphasize the fact that the energy (and the geometry of the molecules in dimer 2) used to calculate the energy of the hydrogen bond linking the molecules of cyclopropylamine is the average energy arising from both molecules.

Parameters	**Experimental**	**Calculated**		
	Solid	*Solid*	*Supercell* (relaxed)	*Dimer 2* (relaxed)
Lattice [Å, °]				
A	5.0741(10)	5.1343	18	5.1343
B	9.7594(10)	10.0127	18	10.0127
C	13.305(2)	13.4640	18	13.4640
α=β=γ	90.0	90.0	90.0	90.0
Z	8	8	1	2
Volume [Å^3]	658.89(19)	692.16	5832	692.16
Space/point group	*Pbca*	*Pbca*	*P1*	*P1*
Geometry [Å, °]				
Dimer 2				
*r*N1…N1	3.166(3)	3.149	—	3.156
*r*N1-H7	0.86(3)	1.035	1.023	1.030
*r*N1…H7	2.30(3)	2.121	—	2.124
∠H7-N1-H6	109(2)	109.4	108.8	108.1
∠H7-N1…N1	167(2)	171.3	—	173.0
Energy [kJ mol^{-1}]				
Sublimation	—	17.2	—	16.2
Total energy [per molecule, eV]	—	-851.87063	-851.69271	-851.86018

Table 3.5: Comparison between the experimental and calculated (PW-DFT) structures for the Phase II crystal structure of cyclopropylamine and the hydrogen bond present in its crystal structure. There are two independent molecules in the dimer 2 calculations and therefore, the total energy reported here for a molecule is the average energy value for both molecules.

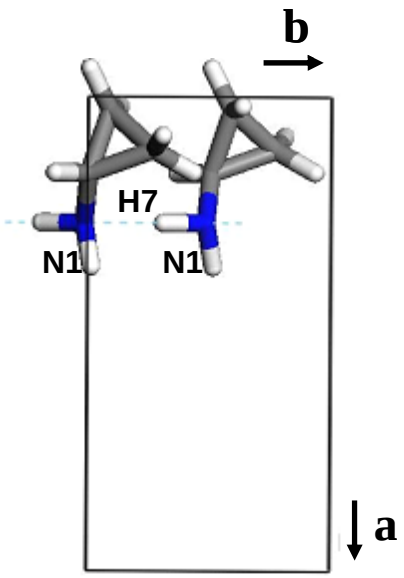


Figure 3.13: View of the model chosen to perform the calculations on the Phase II structure of cyclopropylamine, to mimic the hydrogen bond, dimer 2 [*r*(H7…N1) = 2.285 Å], present in the crystal structure.

3.3.2.3 Calculated energies

All the calculated energies for the crystal structure of Phase II can be found in Table 3.6. A correction was applied to the output CASTEP calculated values due to the application of external pressure, of 1.2 GPa, on the system. Thus, the equation:

$$H = U + PV\text{-}TS \qquad (3.1)$$

is used to correct the enthalpy, H, which is obtained from the PW-DFT calculation by using the external pressure, P, and the volume of the system, V. Since the calculations are carried out at 0 K, the entropy term can be neglected (TS = 0). Thus, the internal enthalpy, U, can be obtained.

The calculated sublimation energy was found to be 4.1 kcal mol^{-1} (17.2 kJ mol^{-1}), which is of a similar magnitude to that estimated for Phase I. This helps lend credibility to the Phase I result, although a closer sublimation energy value to that found for Phase I was expected for Phase II. By definition, the sublimation energy is

the energy of the intermolecular interactions, which need to be broken in order to take the crystal structure from the solid state to the gas phase and therefore, it is expected that the single N-H…N hydrogen bond present in the Phase II crystal structure of cyclopropylamine exhibits an energy of around 4 kcal mol^{-1} (17 kJ mol^{-1}). Thus, we calculate the sublimation energy for Phase I to be higher than the energy for Phase II. Although these calculations refer to the equilibrium geometries and do not take entropy effects into account, it is consistent with experimental observations: we believe that the Phase II structure is only accessible through the application of pressure, and is therefore inherently a higher energy structure.

Model			CASTEP	
	r(D-H…A)	∠(D-H…A)	Energy/ eV	Energy/ kJ mol^{-1}
	r (D…A)	∠(C-D…A)		
Dimer 2	2.30(3)	167(2)		
Periodic	3.166(3)	117.5(1)	-851.86018	16.2
Aperiodic			-851.76666	14.3
Monomer	—	—	-851.69271	—

Table 3.6: Energies obtained per molecule for the Phase II dimer models of cyclopropylamine, and the resulting energies per hydrogen bond (kJ mol^{-1}) (*r*/Å, ∠/°, D = donor, A = acceptor, C = Carbon atom).

The calculated energies for the N1-H…N1 bond were found, as expected, to be 3.9 kcal mol^{-1} (16.2 kJ mol^{-1}) and 3.4 kcal mol^{-1} (14.3 kJ mol^{-1}) for dimer 2 exhibiting periodicity or lacking it, respectively. Thus, the σ-bond cooperativity effect strengthens the hydrogen bond by 11.7%; this is a 0.5 kcal mol^{-1} (1.9 kJ mol^{-1}) energy increase. This, to a certain extent, agrees with what was found for dimer 1 in the Phase I structure of cyclopropylamine; that is, the σ-bond cooperativity effect has an

important effect on the hydrogen bond, although the magnitude of the effect on the hydrogen bond of Phase II is halved.

3.3.3 Computational

3.3.3.1 Crystal structure calculations

Calculations on the Phase I and Phase II crystal structures of cyclopropylamine were performed using the CASTEP package (2002) available through Materials Studio suite of software (Accelerys Inc., 2002). The valence electrons were modelled using a plane-wave basis set expressed at an energy cut-off of 380 eV, which was found to converge the total energies to better than 2.0 meV per unit cell. The electronic core wave function was described using the standard ultrasoft pseudopotentials available with the software package. The symmetry-reduced k-point sets used to sample reciprocal space were generated using Monkhurst-Pack grids (Monkhorst *et al.*, 1977) (dimensions 4 x 2 x 2), giving 8 k-points in the symmetry-reduced first Brillouin zone. The generalised gradient functional PBE was used to model the electronic exchange and correlation.

Although geometry optimisation for the full crystal structure was only performed for the Phase II polymorph of cyclopropylamine, all the atom geometry optimisation calculations presented for the different models (dimer 1, dimer 2 and trimer 1) present in the Phase I and Phase II crystal structures were carried out using the same specifications.

The initial structure of the Phase II polymorph used to perform the geometry optimisation of the structure was taken from the experimental work presented in the first part of this chapter (Lozano-Casal *et al.*, 2005). The optimisation of the atomic positions and unit cell parameters were performed with respect to an external force of 1.2 GPa on alternate cycles using the BFGS method until the convergence criteria were met (maximum energy change per atom = 1×10^{-5} eV, maximum root-mean-square force = 0.03 eV $Å^{-1}$, maximum RMS stress = 0.05 GPa and maximum RMS displacement = 0.001 Å).

3.3.3.2 Supercell calculations

These were performed on single, effectively isolated molecules of cyclopropylamine, for comparison with the energy per molecule in the solid state, and thus allowing the deduction of the total intermolecular interaction energies in the two crystal structures. Zero interaction between the nearest neighbouring cells was obtained by increasing the cell size and observing the change in the total energy. Thus, a cell size of 18 x 18 x 18 $Å^3$ was found to break all intermolecular interactions, without giving rise to overly long computational times. Apart from a reduction in the number of k-points (to one, the gamma point), all other specifications for these calculations were the same as those applied for the corresponding crystal structure optimisation. The energy of the optimised supercell was compared to the solid state energy of the optimised structure to obtain an estimate of the sublimation energy of the system.

3.3.3.3 Dimer and trimer models

These models were constructed to estimate the individual contributions of each of the intermolecular interactions towards the sublimation energy. All calculations performed were geometry optimisations of the atom positions using the same basis set cut-off and convergence criteria as reported for the full Phase II crystal structure optimisation; the k-point sampling grid was reduced to the gamma point.

All the calculations carried out during this computational study were rigorously tested to ensure that the basis sets had reached an acceptable level of convergence. Consequently, the comparison of energies between the supercell and the rest of the calculations (*i.e.*, crystal structure, dimer and trimer calculations) is legitimate: we estimate any error incurred due to inconsistencies in basis sets to be within 0.2 – 0.07 kcal mol^{-1} (i.e., 0.01 – 0.003 eV) for the sublimation and hydrogen bond energies.

3.4 Conclusions

We have determined the high-pressure crystal structure of cyclopropylamine at 1.2 GPa. All the atomic positions were found, including those of the H atoms. Therefore, as the H-atom positions were determined completely, it can be demonstrated directly, and without completely resorting to purely geometrical considerations, that only one of the H atoms of the amino group is involved in the hydrogen bonding. As both H atoms appear to be involved in the hydrogen bonding in the crystal structure of Phase I, this represents a significant difference between the two polymorphs. This change in bonding was also observed in the differences between the Hirshfeld surfaces and fingerprint plots for the two different polymorphs of cyclopropylamine. Finally, the molecular packing environments were studied for the low temperature and the high-pressure Phases I and II of cyclopropylamine and compared with that for perfect body-centred cubic. The molecular environment at 1.2 GPa was found to be less distorted from ideal body-centred cubic packing than that at low temperature. In addition to this, the same conformation of the molecule was found in both Phase I and Phase II structures, as it can be seen from the Hirshfeld surfaces.

The energetics of the Phase I and Phase II crystal structures of cyclopropylamine have been investigated to find that the calculated sublimation energies of the two polymorphs are of 5.2 and 4.1 kcal mol^{-1} (21.9 and 17.2 kJ mol^{-1}) for the ambient and high pressure structures, respectively. This energy difference may be due to the different conditions under which the two different polymorph structures exist (low temperature and high pressure). Both structures present a short N-H…N hydrogen bond, which is involved in the formation of zig-zag molecular chains and exhibit very similar energies of 3.7 and 3.9 kcal mol^{-1} (15.5 and 16.2 kJ mol^{-1}) at ambient and high pressure conditions, respectively. The calculated energy for the weak N-H…N contact present in the ambient pressure structure, Phase I, of cyclopropylamine is of 1.5 kcal mol^{-1} (6.4 kJ mol^{-1}), which is almost a third of the energy of the hydrogen bond, which forms zig-zag chains. This shows that the energy of a hydrogen bond is correlated to the geometry of the interaction (i.e., H…A distance, D-H…A angle).

Finally, the strength of the σ-bond cooperativity effect on the intermolecular interactions present in the crystal structures of the Phase I and Phase II polymorphs of cyclopropylamine has been described. Results show that this effect strengthens the dimer interactions, both in the Phase I and Phase II crystal structures, by about 26 and 12%, respectively. Thus, the presence of the σ-bond cooperativity effect on the dimer interactions is relevant to their formation and stabilisation.

3.5 References

Allan, D. R., Clark, S. J. (1999a). *J. Phys. Rev. B,* **60**, 6328.

Allan, D. R., Clark, S. J. *J.* (1999b). *Phys. Rev. Lett.* **82**, 3464.

Allan, D. R., Clark, S. J., Brugmans, M. J. P., Ackland, G. J., Vos, W. L. (1998). *Phys. Rev. B,* **58**, R11809.

Allan, D. R., Clark, S. J., Dawson, A., McGregor, P. A., Parsons, S. (2002). *J. Chem. Soc.* **8**, 1867.

Allan, D. R., Clark, S. J., Ibberson, R. M., Parsons, S., Pulham, C. R., Sawyer, L. (1999c). *Chem. Commun.* **8**, 751.

Allan, D. R., Clark, S. J., Parsons, S., Ruf, M. (2000). *J. Phys. Condens. Matter*, **12**, L613.

Allan, D. R., Parsons, S., Teat, S. J. (2001). *J. Synchr.Rad.* **8**,10.

Angel, R. J. *J.* (2004). *Appl. Cryst.* **37**, 486.

Berstein, J., Davis, R. E., Shimoni, L., Chang, N. L. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555.

Blatov, V. A. (2004). *Crystallogr. Rev.* **10**, 249-318.

Blatov, V. A., Shevchenko, A. P., Serezhkin, V. N. (1999). *Russ. J. Coord. Chem.* **25**, 453-465.

Blessing, R. H. (1995). *Acta Cryst.* **A51**, 33.

Blessing, R. H. (1997). *J. Appl. Cryst.* **30**, 421.

Boer, J. S. A. M.; Schenk, H.; Stam, C. H. (1986). *Rec. Trav. Chim.* **105**, 434.

Brock, C. P & Duncan, L. L. (1994). *Chem. Mater*. **6**, 1307.

Bruker-AXS (1997-2001). *SMART*, Version 5.049-5.059. Bruker-AXS, Madison, Wisconsin, USA.

Bruker-AXS (1999). *GEMINI*, Version 1.01. Bruker-AXS, Madison, Wisconsin, USA.

Bruker-AXS (2002). *SAINT*, Version 6. Bruker-AXS, Madison, Wisconsin, USA.

Dawson, A., Allan, D. R., Parsons, S., Ruf, M. (2004). *J. Appl. Cryst.* **37**, 410.

Desiraju, G. & Steiner, T. (2001). *The Weak Hydrogen Bond*. Oxford University Press, USA.

Fabbiani, F. P. A., Allan, D. R., Dawson, A., David, W. I. F., McGregor, P. A., Oswald, I. D. H., Parsons, S., Pulham, C. R. (2003). *Chem. Comm.* **24**, 3004.

Grimwood, D.; Wolff, S. K.; McKinnon, J.; Spackman, M.; Jayatilaka, D. (2004). *Crystal Explorer version 1.0.3*, University of Western Australia, Australia.

Hirshfeld, F. L. (1977). *Theor. Chim. Acta*. **44**, 129.

Jönsson, P. G. (1976). *Acta Cryst.* **B32**, 232.

Lozano-Casal, P. ; Allan, D. R.; Parsons, S. (2005). *Acta Cryst.* **B61**, 717-723.

Materials Studio (2001-2005). Accelerys Software Inc.

McGregor, P. A.; Allan, D. R.; Parsons, S.; Pulham, C. R. (2005). *Acta Cryst*. **B61**, 449-454.

McKinnon, J. J.; Spackman, M. A.; Mitchell, A. S. (2004). *Acta Cryst.* **B60**, 627.

Merrill, L & Basset, W. *A.* (1974). *Rev. Sci. Instrum.*, **45**,290.

Monkhorst, H. J. & Pack, J. D. (1977). *Phys. Rev. B*, **13**, 5188-5192.

Morrison, C. A. & Siddick, M. M. (2003). *Chem. Eur. J.* **9**, 628-634.

Narten, A. H. & Habenschuss, A. (1984). *J. Chem. Phys.* **80**, 3387.

Oswald, I. D. H.; Allan, D. R.; Motherwell, S. W. D.; Parsons, S. (2005). *Acta Cryst.* **B61**, 69-79.

Peresypkina, E. V.& Blatov, V. A. (2000). *Acta Cryst.* **B56**, 501.

Segall, M. D.; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. (2002). *J. Phys. Condens. Matter*, **14**, 2717.

Sheldrick, G. M. (1997a). *SHELXTL*. Bruker-AXS, Madison, Wisconsin, USA.
Sheldrick, G. M. (1997b). *SHELXS97*. University of Göttingen, Germany.

Sheldrick, G. M. (1997c). *SHELXL97*. University of Göttingen, Germany.

Shipman, L. L.; Burgess, A. W.; Scheraga, H. A. (1976). *J. Phys. Chem.* **80**, 52.

Steiner, T. (2002). *Angew. Chem. Int. Ed.* **41**, 48.

Tauer, K. J. & Lipscomb, W. N. (1952). *Acta Cryst.* **5**, 606.

Torrie, B. H.; Weng, S. X.; Powell, B. M. (1989). *Molec. Phys.* **67**, 575-81.

Chapter 4

Computational Study of α-Glycine

4.1 Introduction

Amino acids are the building blocks of proteins, which in turn form part of every living being from bacteria to humans. In the condensed phases amino acids exist in the charged zwitterionic state, whereas it is the neutral form which is more stable in the gas phase (Császár & Perczel, 1999). The possibility of forming a strong hydrogen bonding network, dominated by electrostatic interactions between charged molecules, is the driving force that dictates the crystal packing arrangement.

The computational study that we report here is an integral part of a high-pressure X-ray diffraction experimental study, aimed at investigating the behaviour of the interactions between amino acid zwitterionic molecules (Dawson *et al.*, 2005; Moggach *et al.*, 2005). In this work amino acids have been chosen for investigation as they are small organic molecules, suitable for accurate experimental and computational study, and can act as model systems to begin to understand how much bigger and more complex systems like proteins behave under high-pressure conditions. In such work calculations can offer valuable insight to guide and complete the high-pressure crystallographic experiments, as the Merrill-Bassett diamond anvil cell (Merrill & Bassett, 1974) used to compress the sample restricts the amount of reciprocal (diffraction) space that can be successfully harvested. Consequently, reported structures are typically of lower precision compared to their ambient pressure counterparts and hydrogen atoms have to be placed by hand. It has been demonstrated that quantum mechanical simulations can successfully complete these partial experimental structures (Allan *et al.*, 1999; Walker *et al.*, 2005). Additional important information can also be obtained from the calculations performed, such as lattice and sublimation energies, energies of hydrogen bonding and proton transfer energies. The lattice energy is equal to the amount of energy that has to be supplied to the crystal structure lattice to allow the molecules to be separated to infinity whilst still retaining the same geometry. It is therefore a direct measure of the average hydrogen bond energy present in the lattice. In our calculations we obtain the lattice energy by comparing the energy of one molecule in the crystal structure with that of one molecule in the absence of its neighbours. The lattice energy is then converted to the sublimation energy, which is the energy that must be applied to drive the molecule

from the solid state to the gas phase, by allowing the isolated molecular geometry coordinates to optimise. Finally, the difference between the sublimation and the lattice energies yields the proton transfer energy; this is the energy that must be applied to the molecule to transfer a proton from the amino group (zwitterion) to the carboxylic group (neutral).

In this study we want to partition the total interaction energy into the different hydrogen bonds present in the structure. In theory there is no reason why we cannot do this: we can build models to calculate the individual energies of the interactions. However, these models will contain more than just the hydrogen-bond linkage under investigation. Secondary interactions, that is short-range cross-interactions between other frontier atoms of each molecule involved in the hydrogen bond formation will also be present, and cannot be formally separated from the hydrogen bond energy. A further complication for amino acids is that when two zwitterionic states are taken out of the solid state and optimised as an isolated dimer, the energy minimum obtained corresponds to the proton-transferred neutral form. For this reason, all the calculations run to obtain the individual contributions to the total intermolecular interaction energy were performed as single-point energy calculations. In fact, this complication actually simplifies matters, as we would expect the bulk of the interaction energy for charged molecules to be described by simple electrostatics (*i.e.* the Coulombic potential), which we can estimate based on calculated Mulliken charges. We can therefore break down the total electrostatic interaction energy to individual atom-atom electrostatic pair contributions, and in this way separate out an estimation of the (electrostatic) hydrogen bond energy from the secondary interactions.

In this study we have applied several different computational approaches in order to estimate the strengths of hydrogen bond interactions in α-glycine, which was chosen as there exists a well of literature values that can be used as references to test the computational procedure. Initial calculations were performed with the CASTEP code (Segall *et al.*, 2002), using plane-wave density functional theory (PW-DFT), which makes use of periodic boundary conditions and thus allows us to estimate the lattice, sublimation and proton transfer energies, in addition to the total interaction energies for the different hydrogen bonds present in the amino acid crystal structure

studied. In order to benchmark the interaction energies obtained by this method, a second series of calculations on the same model systems (but minus the periodic boundary conditions) was undertaken *via* the more conventional localised basis set approach using GAUSSIAN 98 (Frisch *et al.*, 1998). The electrostatic contributions to the total interaction energies was then obtained by performing Ewald summations using the program GULP (Gale *et al.*, 2003), taking the calculated geometries and Mulliken charges obtained by PW-DFT as input. Finally, an in-house program, V_Coulombic, evaluated the electrostatic contributions from the individual atom-atom pair interactions present.

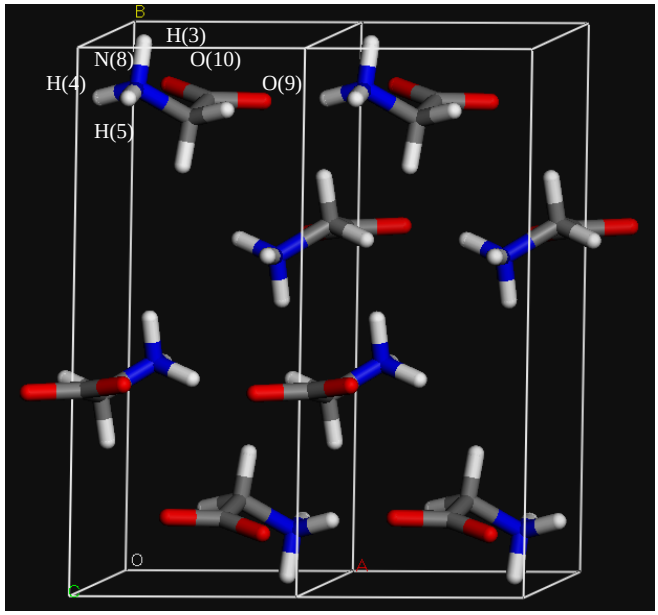
4.2 Results and discussion

The principal aims of this work are to demonstrate that PW-DFT calculations can achieve reliable estimates for sublimation and lattice enthalpies, and energies of individual hydrogen bonds present in the crystal structures of amino acids. Before we present results obtained for the test case, α-glycine, it is essential to first summarise the literature currently available for this simple amino acid system.

4.4.1 Crystal structure of α-glycine at ambient conditions

The starting model to carry out a full quantum mechanical geometry optimisation for α-glycine was taken from the single-crystal X-ray diffraction study at ambient conditions performed by Drebushchak *et al.* (2002). The crystal structure is characterised by double sheets of molecules in the charged zwitterionic state linked by a network of relatively strong electrostatic $N^{+} - H \ldots {}^{-}O{=}C$ hydrogen bonds, which binds molecules in layers along the 'a-c' face. Pairs of sheets, connected two by two by weaker bonds (electrostatic to Van der Waals interactions), lie anti-parallel with respect to each other to form double layers (Marsh *et al.*, 1958; Jönsson *et al.*, 1972; Power *et al.*, 1976; Day *et al.*, 2003; Dawson *et al.*, 2005) [Figure 4.1 (a)].

(a)



(b)

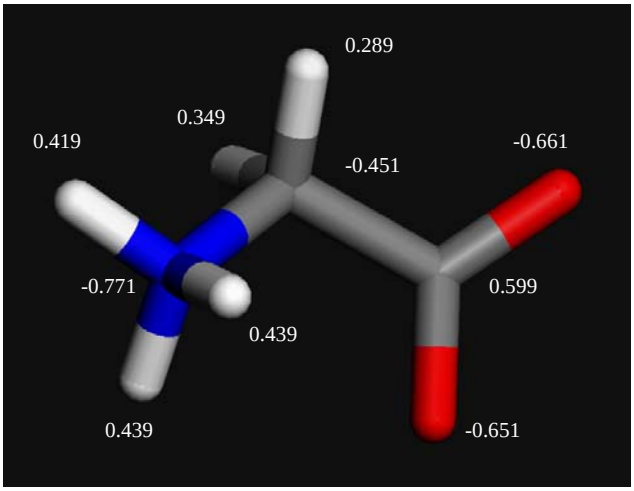


Figure 4.1: (a) The crystal structure and atom labelling of α-glycine. (b) Calculated Mulliken charges for the atoms forming the α-glycine molecule in the solid state structure.

4.4.2 Literature energies

The α-glycine molecule has been extensively explored in terms of conformational studies in the gas phase (Yu *et al.*, 1995; Császár *et al.*, 1999), polymorphism (Chisholm *et al.*, 2005), proton transfer (Takagi *et al.*, 1958; Voogd *et al.*, 1981; Zhang *et al.*, 1999) and general thermodynamic properties (No *et al.*, 1990, 1994; Bisker-Leib *et al.*, 2002; Momany *et al.*, 1974; Derissen *et al.*, 1977; Svec *et al.*, 1965; Gaffney *et al.*, 1977), which are summarised in the compendium of sublimation enthalpies published by Chickos *et al.* (2002).

A summary of the relevant values for the sublimation and lattice enthalpies and the proton transfer energy is shown in Table 4.1. From this it can be seen that for the lattice enthalpy a wide range in calculated values has been reported (58-103 kcal mol^{-1} or 242-430 kJ mol^{-1}). For the sublimation enthalpy there are important differences in the reported experimental values which depend on the way they have been measured, with values obtained by proton transfer mass spectra of the order 23 kcal mol^{-1} (96 kJ mol^{-1}), and those by the Knudsen cell effusion method around 32.6 kcal mol^{-1} (136.3 kJ mol^{-1}).The reason for this difference was explained by No *et al.* (1994). Despite the differences in the literature values for the sublimation and lattice enthalpies, the reported values for the proton transfer energies are all largely consistent, falling between 34.3 and 37.3 kcal mol^{-1} (143.4 and 155.9 kJ mol^{-1}).

No.	Experimental method	Computational method	$\Delta Hsub_{exp}$	$Elat(0)_{calc}$	$Ept(0)_{calc}$	Reference
1	Knudsen cell effusion	SCF/6-31G*	32.6	68	34.32	No *et al.*,1994
	Proton transfer mass spectra	SCF/6-31G*	23	58	-	
2	-	SCF/6-31G**	-	64.8	30.78	Kwon *et al.*,1995
	-	-	-	43.1 (electrostatic contribution only)	-	Kwon *et al.*, 1996
3	-	SCF/DZP	-	49.2 (electrostatic contribution only)	37.32	Voogd *et al.*, 1981
4	-	RHF/4-31G*	-	-	29	Tse *et al.*, 1978
5	-	SCF/6-31G**	34.7	73.2	-	Boek *et al.*, 1991
6	-	CNDO/2	-	26	-	Momany *et al.*, 1974
7	-	Free energy of hydratation	-	103	-	Shimura *et al.*, 1951
8	Knudsen cell effusion	-	32.6(1)	-	-	Svec *et al.*, 1965
9	Proton transfer mass spectra	-	23(1)	-	-	Gaffney *et al.*, 1975
10	Knudsen cell effusion	-	32.0(8)	-	-	Yu *et al.*, 1995

Table 4.1: Summary of the literature sublimation, lattice and proton transfer energies available for α-glycine (kcal mol^{-1}).

4.4.3 Geometry optimisation

The results obtained by PW-DFT calculations for the geometry optimisation of the α-glycine structure can be found in Table 4.2, together with the experimental data for comparison. From this it can be seen that in general the computational results are in good agreement with the experimental structure.

The differences in cell parameters are 1.9, 5.6 and – 0.3 % for the 'a', 'b' and 'c' directions respectively, and the change in the β angle is of the order of 2.0%. The cell volume increased by 8.9% compared to the experimental value; some volume expansion is always expected in a DFT simulation since the dispersion interaction energies are neglected. The simulation reported heavy-atom positions largely consistent with the input experimental ones, with differences in the bonding parameters of the order of picometers [*e.g.* $r(\mathrm{C-N})_{\mathrm{exp}} = 1.479$ Å, $r(\mathrm{C-N})_{\mathrm{calc}} = 1.473$ Å]. As expected, the main differences observed relate to the hydrogen-atom positions. The X-ray diffraction experiment places the atoms in positions corresponding to electron density maxima. Thus, for the lightest atom, hydrogen, there is an inherent uncertainty in its position due to its low electron density (only one valence electron). In contrast, the definition of bond length in the simulation relates to the inter-nuclear distance; in all likelihood the hydrogen atom positions are more accurately known from the quantum mechanical geometry optimisation.

Parameters	**Experimental[a]**		**Calculated**	
	Solid	Solid	Supercell (not relaxed)	Supercell (relaxed)
Lattice [Å, °]				
A	5.106(1)	5.202	16	16
B	11.979(5)	12.647	16	16
C	5.463(2)	5.447	16	16
β	111.75(2)	109.51	109.51	109.51
Z	4	4	1	1
Volume [$Å^3$]	310.4	337.8	3860.7	3860.7
Space/point group	*P21/n*	*P21/n*	*P1*	*P1*
Geometry [Å, °]				
Dimer 1				
rN8…O9	3.081	2.934	—	—
rN8-H3	0.941	1.044	1.044	—
rO9…H3	2.197	1.931	—	—
∠H3-N8-H5	107.1	107.6	107.6	—
∠H3-N8…O9	156.1	160.1	—	—
Dimer 2				
rN8…O9	2.852	2.819	—	—
rN8-H4	0.949	1.054	1.054	1.021
rO9…H4	1.917	1.776	—	—
∠H3-N8-H4	108.4	107.0	107.0	108.9
∠H4-N8…O9	167.8	169.3	—	—
Dimer 3				
rN8…O10	2.771	2.722	—	—
rN8-H5	1.016	1.071	1.071	1.02
rO10…H5	1.770	1.652	—	—
∠H4-N8-H5	109.6	109.0	109.0	109.9
∠H5-N8…O10	167.9	176.7	—	—
Energy [kcal mol^{-1}]				
Lattice	58-103[b]	—	57.1	—
Sublimation	23-34.7[b]	—	—	26.1
Proton Transfer	34.3-37.3[b]	31.0	—	—
Total energy [per molecule, eV]	—	-1541.59162	-1539.11988	-1540.46040

[a] From Drebushchak *et al.* (2002) (experimental data were collected at ambient conditions).
[b] See Table 1 for appropriate references.

Table 4.2: Comparison between the experimental and calculated (PW-DFT) structures for the crystal structure of α-glycine.

4.4.4 Calculated energies

The calculated lattice energy was found to be 57.1 kcal mol^{-1} (238.7 kJ mol^{-1}), which lies just outside the lower range of values reported for this property (58-103 kcal mol^{-1} or 242-430 kJ mol^{-1}). The calculated sublimation energy, at 26.1 kcal mol^{-1} (109.1 kJ mol^{-1}), is in better agreement with the reported range (23-34.7 kcal mol^{-1} or 96-145.0 kJ mol^{-1}), particularly those obtained by proton transfer mass spectra (*ca.* 23 kcal mol^{-1} or 96 kJ mol^{-1}). Any discrepancy between calculation and experiment is likely to be attributed to two reasons: firstly, we would expect the calculation to underestimate the interaction energy as van der Waals (dispersion) interactions have been neglected. Secondly, whereas the experiments are performed at elevated temperatures, our calculations are performed at 0 K and neglect the zero-point energy correction. That said, the calculated values are about the right order of magnitude, and the proton transfer energy, which is the difference between the lattice and sublimation energies, at 31.0 kcal mol^{-1} (129.6 kJ mol^{-1}), is close to the range reported (34.3-37.3 kcal mol^{-1} or 143.4-155.9 kJ mol^{-1}).

In order to further benchmark the thermodynamic results obtained by PW-DFT, and to estimate the electrostatic contribution to the lattice energy, an Ewald summation was performed using the program GULP, taking the optimised structure and calculated Mulliken charges obtained by the quantum mechanical simulation [Figure 4.1 (b)] as input. The value thus obtained for α-glycine was 58.6 kcal mol^{-1} (244.9 kJ mol^{-1}), which is in very close agreement with the total energy result obtained by PW-DFT (57.1 kcal mol^{-1} or 238.7 kJ mol^{-1}). From this it is verified that the bulk of the lattice energy is attributed to an electrostatic component, which was to be expected from the zwitterionic nature of the amino acid molecules making up the crystal lattice.

One of the principal aims of this work is to assign energy values to the individual hydrogen bond interactions present in the crystal lattice. It is generally accepted that the α-glycine crystal structure is formed by three different hydrogen bonds; we will therefore have to construct three different models (labelled dimer models 1-3, see Figure 4.2 and Table 4.2) to obtain each interaction in isolation. To

take each model in turn, dimer 1 concerns $N_8 - H_3 \ldots O_9$ (1.931 Å, 160.1°). The interaction is aperiodic (*i.e.* it lacks σ-bond cooperativity), lies along the unit cell b-direction and forms rings labelled $R_2^2(10)$ in graph set notation (Bernstein *et al.*, 1995). Dimer 2 describes the periodic hydrogen bond $N_8 - H_4 \ldots O_9$ (1.776 Å, 169.3°), which forms a zig-zag chain $C(2)$ along the a-direction in the unit cell. Dimer 3 relates to the periodic $N_8 - H_5 \ldots O_{10}$ hydrogen bond (1.652 Å, 176.7°), which links two molecules into zig-zag chains along the unit cell c-axis [graph set notation $C(2)$].

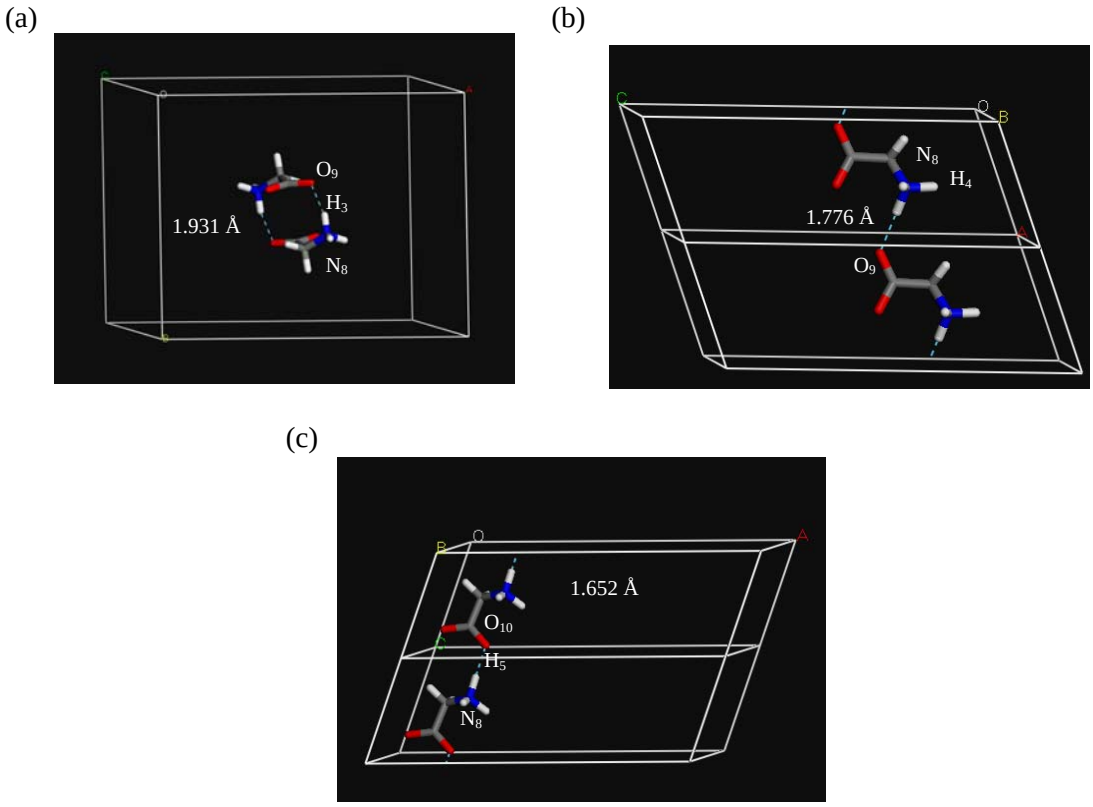


Figure 4.2: The three dimer models used in the calculation of the individual hydrogen bond energies (dimers 1, 2 and 3, respectively). Note the bond order in each model is 1.

On the basis of geometries alone, we would expect the interaction present in dimer 1 to be the weakest, and dimer 3 the strongest (Jeffrey *et al.*, 1997). At face value the bond strengths obtained by all the modelling techniques (see Table 4.3) do not fit in with the expected textbook trends. The PW-DFT values, at 22.7, 12.8 and

29.8 kcal mol^{-1} (94.9, 53.8 and 124.6 kJ mol^{-1}) for models 1-3, respectively, sum to 65.4 kcal mol^{-1} (273.4 kJ mol^{-1}). This is very close to the calculated lattice energy (57.1 kcal mol^{-1} or 238.7 kJ mol^{-1}), which lends assurance to the accuracy of the individual interaction energy calculations. Moreover, since it appears that we can calculate the energies of the three hydrogen bonds separately with some degree of confidence, it suggests that the level of correlation between the bonds must be low. At face value, however, the bond strengths obtained by all the modelling techniques (see Table 4.3) do not fit in with the expected 'textbook' trends [Jeffrey *et al.*, 1997]. The interaction present in dimer 1 is considerably stronger than that in dimer 2, and the values obtained fall over a very wide range, both points necessitating further investigation.

In search of an explanation we studied the electrostatic contribution to the hydrogen bond interactions using an in-house program V_Coulombic, which simply calculates the Coulombic interaction potential for individual atom-atom pairs from a given set of coordinates and point charges. Using the same input dimer models as described above and summing the individual Coulombic energy contributions for the aperiodic dimer models 1 to 3 results in 20.5, 3.0 and 28.4 kcal mol^{-1} (85.7, 12.5 and 118.7 kJ mol^{-1}), respectively, which is in good agreement with those obtained by a full Ewald sum for the same input models provided by the program GULP and that obtained by the full quantum mechanical treatment by CASTEP (see Table 4.3). We can therefore confidently conclude that the same unusual trend in total interaction strengths is manifest in just the electrostatic component of the dimer interactions.

Upon closer examination of just the H…O atom pair making up the three different hydrogen bonds the electrostatic interaction energy is 49.9, 54.3 and 54.9 kcal mol^{-1} (208.6, 227.0 and 229.5 kJ mol^{-1}) for dimer models 1 to 3 respectively. The trend in these numbers is now in agreement with the expected strengths of the different interactions in terms of their donor-H acceptor geometries (Jeffrey, 1997); that is the hydrogen bond expected to be the weakest (dimer 1) is now correctly assigned the lowest calculated electrostatic interaction energy.

Model			V_Coulombic	CASTEP		GULP		GAUSSIAN 98		
	r(D-H…A) *r* (D…A)	∠(D-H…A) ∠(C-D…A)	Energy/ H- bond	Energy/ eV	Energy/ H- bond	Energy/ eV	Energy/ H- bond	Energy/ Hartrees	Energy/ H- bond This work	 Literature[a]
Dimer 1	1.931	160.1								
Aperiodic	2.934	97.4	20.5	-1540.10229	22.7	0.89032	20.6	-284.489695	24.3	25.8
Dimer 2	1.776	169.3								
Periodic	2.819	116.7	—	-1539.20072	12.8	0.37990	8.8		—	—
Aperiodic	2.819	116.7	3.0	-1538.93916	4.1	0.10961	2.5	-284.454462	6.8	3.3
Dimer 3	1.652	176.7								
Periodic	2.722	112.5	—	-1540.411974	29.8	1.48296	34.2		—	—
Aperiodic	2.722	112.5	28.4	-1539.643421	24.2	1.32972	30.7	-284.468103	23.8	22.5
TOTAL										
Periodic					54.4		63.6			
Aperiodic			51.9		38.6		53.8		54.9	51.6
Monomer	—	—	—	-1539.119884	—	—	—	-284.446641		—

[a] (Volkov *et al.*, 2004)

Table 4.3: Energies obtained per molecule for the α-glycine dimer and monomer models, and the resulting energies per hydrogen bond (kcal mol^{-1}) (D – H … A, C-D…A; *r*/Å, ∠/°, D = donor, A = acceptor, C = Carbon atom).

The importance of the contributions from the individual electrostatic atom-atom pair interactions is also demonstrated by providing an explanation for the wide differences present between the different dimer models. The output from V_Coulombic indicates that the electrostatic energies arising from the different atom-atom pair interactions (*i.e.* secondary interactions) fall in a very wide range of positive (repulsive) and negative (attractive) energy values (see Figure 4.3). This finding was also reported by Popelier *et al.* (2002) when studying the stability of DNA base pairing. Thus the huge variation in interaction energy is simply due to the difference in conformation of the molecules in the two different models. In dimer 2 the molecules are arranged in an 'anti' conformation, [*i.e.* O_9 is 'anti' with respect to N_8, forming an O_9-C_6-C_7-N_8 angle of 149.4°] whereas in dimer 3, the conformation is 'syn' [that is O_{10} and N_8 are in a "syn" conformation forming a O_{10}-C_6-C_7-N_8 angle of 31.2°]. This results in important differences in the atom-atom electrostatic interactions between molecules, *e.g.* the N_8…O_9 intermolecular interaction in dimer 2 is 20.0 kcal mol^{-1} (83.6 kJ mol^{-1}), whereas in dimer 3 it represents an energy of 195.5 kcal mol^{-1} (817.2 kJ mol^{-1}). Similarly, the N_8…O_{10} interaction is 24.9 kcal mol^{-1} (104.1 kJ mol^{-1}) in dimer 2 and 61.3 kcal mol^{-1} (263.8 kJ mol^{-1}) in dimer 3 (Figure 4.3). Additionally, when comparing the overall repulsive and attractive energies arising from the atom-atom interactions between molecules in dimer models, it was found that the three interactions present repulsive energies of 1052.6, 847.9 and 822.5 kcal mol^{-1} (4399.9, 3544.2 and 3438.0 kJ mol^{-1}) for dimers 1 to 3, respectively, and attractive energies of -1093.0, -850.8 and -851.0 kcal mol^{-1} (4868.7, 3556.3 and 3557.2 kJ mol^{-1}). In the case of dimer 1 the magnitude of the repulsive and attractive energies is considerably larger than those for dimers 2 and 3. This is likely to be due to the fact that in dimer 1 the molecules are arranged in a cyclic conformation, whereas dimers 2 and 3 are open chains. As can be seen from the numbers cited above, while the repulsive energy is higher for dimer 2 than for number 3 the attractive energy does not seem to be influenced significantly by the "anti" and "syn" conformations of the molecules.

Finally, for completeness we also report on the interaction strengths obtained for the three different dimer models by *ab initio* molecular orbital theory. Volkov *et al.* (2004) reported similar values (of 25.8, 3.3 and 22.5 kcal mol^{-1} or 107.8, 13.8 and

94.0 kJ mol^{-1}), for aperiodic models 1-3, respectively) obtained from calculations performed on glycine dimers in the gas phase. These values were also largely consistent with calculations we performed ourselves on the same model systems using the GAUSSIAN98 simulation package (see Table 4.3).

The σ-bond cooperativity effect is more significant in dimer model 2 than in model dimer 3. Nevertheless the energies of the hydrogen bonds present in both dimer models strengthen when this effect is present. Thus, the results obtained from CASTEP show that the hydrogen bond in dimer 2 exhibits an energy value of 12.8 kcal mol^{-1} (7.9 kJ mol^{-1}) when periodic and 4.1 kcal mol^{-1} (17.1 kJ mol^{-1}) when aperiodic. This implies that the presence of the σ-bond cooperativity effect strengthens the hydrogen bond by a factor of three. When the same calculations were carried out using GULP, the results obtained were relatively similar, of 8.8 and 2.5 kcal mol^{-1} (36.8 and 10.11 kJ mol^{-1}) for the periodic and aperiodic interactions, respectively (Table 4.3). In the case of dimer model 3 σ-bond cooperativity was also demonstrated to be an appreciable effect: CASTEP calculated an energy of 29.8 kcal mol^{-1} (124.6 kJ mol^{-1}) for the periodic hydrogen bond and 24.2 kcal mol^{-1} (101.2 kJ mol^{-1}) for the aperiodic interaction. These results were also largely supported by calculated values from GULP (see Table 4.3).

(a) (b) (c)

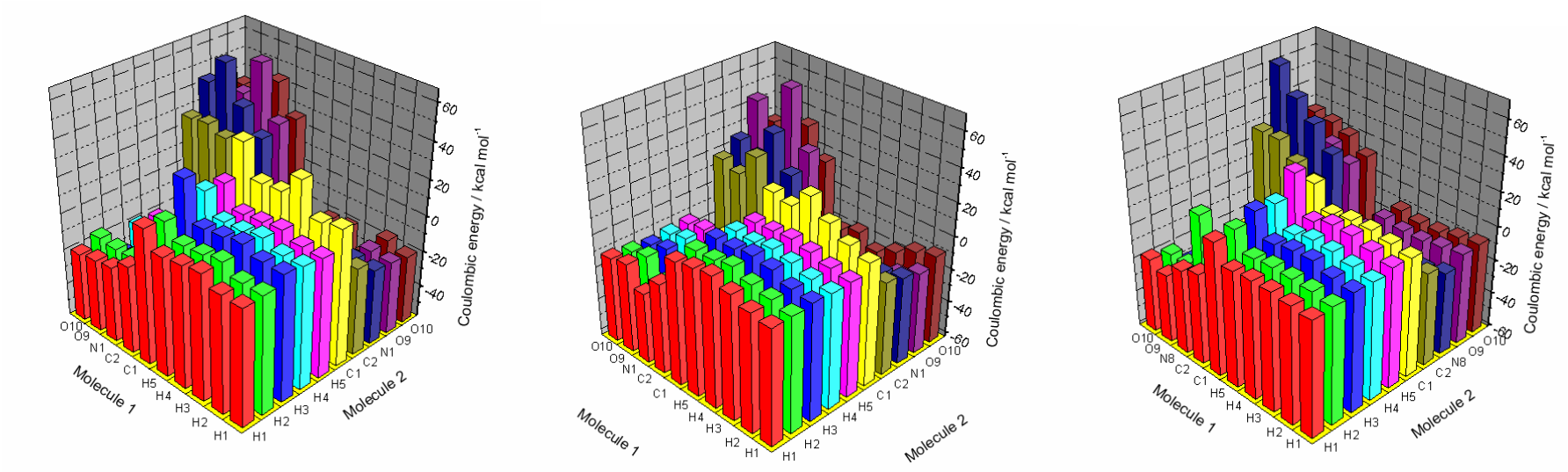


Figure 4.3: Histogram of the atom-atom contributions to the total electrostatic interaction energy (kcal mol^{-1}) marked according to atom type for (a) model 1, (b) model 2 and (3) model 3 in α-glycine.

4.3 Compressibility of the α-glycine crystal structure according to the computational results

The effect of high pressure on the crystal structures of the different glycine polymorphs (i.e. α, β, γ, δ and ε) was investigated and published by Dawson *et al.* (2005). In this work, we are particularly interested in the compressibility study of α-glycine, since it is the test molecule used to prove the computational method. When going from ambient pressure (i.e. 0 GPa) to 6.2 GPa, the unit-cell dimensions of the monoclinic $P2_1/n$ structure decreased by 4.6, 7.0 and 1.6% for the 'a', 'b' and 'c' directions respectively, whereas the 'β' angle increased from 111.740(5)° to 116.888(11)° (Table 4.4). Therefore, the most compressible direction is along the 'b' axis, followed by the 'a' axis and the 'c' axis, which changes only by a small amount. These changes in the unit cell dimensions are a direct consequence of the effect of pressure on the hydrogen bonding present in the crystal structure, due to the alignment of the hydrogen bonds with the three cell vectors. For instance, the N_8-H_5…O_9 hydrogen bond is most likely responsible for the shortening of the c-axis since it forms *C(5)* chains along the 'c' direction. The N_8-H_4…O_9 interaction, however, bridges chains of molecules to form molecular layers parallel to the 'ac' plane, participating in the decrease of length of the 'a' and 'c' axes. Finally, the N_8-H_3…O_{10} interaction forms $R_2^2(10)$ ring motifs, parallel to the 'b' axis, to build up a bilayer structure, causing the compression of the 'b' axis when pressure is applied to the system. In order to justify the changes produced in the unit-cell dimensions, the changes in the three different hydrogen bonds must be investigated. This study was also carried out by Dawson *et al.* (2005), who reported changes in the N…O hydrogen bond distances of 1.4, 4.7 and 1.1% for dimers 1-3, respectively. In this work we want to investigate the correlation between the compressibility of the hydrogen bonds and their energetics; that is, the hydrogen bonds displaying the largest compressions are expected to be weaker than those presenting smaller changes. The computational study carried out on the hydrogen bonding present in the α-glycine crystal structure concluded that the energies of the interactions present in dimers 1-3 are of around 22, 12 and 30 kcal mol^{-1} (92, 50 and 125 kJ mol^{-1}), respectively (see Table 4.3). According to these energy values, the largest compression should happen for those

interactions present in dimer 2, since the energy of the hydrogen bond is only about 12 kcal mol^{-1} (50 kJ mol^{-1}). This correlates with what was found experimentally, since the N_8-H_4…O_9 presents a considerable decrease in length of around 4.7%. The hydrogen bond with the lowest compressibility, N_8-H_5…O_{10}, was modelled in our dimer 3 calculation, and correlates with the highest energy calculated in our study, of around 30 kcal mol^{-1} (125 kJ mol^{-1}). From these results, a clear correlation can be seen between the energy and the compressibility of the hydrogen bonding in the crystal structure of α-glycine (see Table 4.4). Whilst this was pleasing to observe it was expected since pressure is a thermodynamic variable which is widely used in structural studies to tune the properties of the hydrogen bond in order to get a better understanding of its nature. Despite these expectations, little is actually known to quantify the relationship between the compressibility studies of amino acids and the energetics of these systems. Therefore, the work presented here constitutes a first step towards this objective.

Polymorph	**α-Glycine** (Dawson *et al.*, 2005)		
Pressure / GPa	**0**	**6.2**	**Change/%**
Crystal system	Monoclinic	Monoclinic	
Space group	*$P2_1/n$*	*P21/n*	
a/Å	5.1047(3)	4.8690(7)	-4.6
b/Å	11.9720(14)	11.139(3)	-7.5
c/Å	5.4631(3)	5.3777(10)	-1.6
β/°	111.740(5)	116.888(11)	+4.6
V/$Å^3$	310.10(4)	260.14(10)	-16.1
Z	4	4	
Hydrogen bonds	**D(N …O)/Å**	**D(N …O)/Å**	
N8-H5…O9^{iii} (dimer 1)	3.0748(11)	3.031(18)	-1.4
N8-H4…O9^{ii} (dimer 2)	2.8507(10)	2.721(9)	-4.7
N8-H3…O10^{i} (dimer 3)	2.7703(9)	2.740(7)	-1.1

i) x, y, z+1; ii) x-1, y, z; iii) 2-x, 2-y, -z

Table 4.4: Crystallographic Data for α-glycine (Dawson *et al.*, 2005).

4.4 Computational

4.4.1 Quantum mechanical simulations

4.4.1.1 Crystal structure calculations

Calculations on the crystal structure of α-glycine–were performed using the CASTEP 2.2 package (Segall *et al.*, 2002) available through Materials Studio suite of software (Accelerys Inc., 2001-2005). The valence electrons were modelled using a plane-wave basis set expressed at an energy cut-off of 380 eV, which was found to converge the total energies to *ca.* 0.01 eV per atom. The electronic core wave function was described using the standard ultrasoft pseudopotentials available with the software package. The symmetry-reduced k-point sets used to sample reciprocal space were generated using Monkhorst-Pack grids (Monkhorst *et al.*, 1977) (dimensions 4 x 2 x 4), giving 8 k-points in the symmetry-reduced first Brillouin. The generalised gradient functional PBE (Perdew *et al.*, 1996) was used to model the electronic exchange and correlation.

The initial structure used to perform the geometry optimisations was taken from the ambient pressure and temperature X-ray diffraction study by Drebushchak *et al.* (2002). The optimisations of the atomic positions and unit cell parameters were performed on alternate cycles using the BFGS method until the convergence criteria were met (maximum energy change per atom = 1 x 10^{-5} eV, maximum root-mean-square force = 0.03 eV $Å^{-1}$, maximum RMS stress = 0.05 GPa and maximum RMS displacement = 0.001 Å).

A population analysis calculation was performed on the optimised structure to obtain the Mulliken charges for each atom present in the crystal lattice. Since the Mulliken population analysis is formulated in terms of atomic orbitals (Mulliken, 1955), the electronic wave function from a plane wave is first projected to an atomic orbitals basis set using the scheme of Sanchez-Portal (Sanchez-Portal *et al.*, 1995). Given the atomic basis, its quality is evaluated by its projection into the plane-wave

eigenfunctions, then it is optimised by maximising its projection and the associated tight-binding Hamiltonian and energy bands are obtained; finally, the population analysis is performed in a natural way The charge spilling parameter for the spin component, which measures how much of the subspace of the Hamiltonian eigenstates falls outside the subspace spanned by the atomic basis, was found to be 1.22, indicating a successful charge partitioning was obtained.

4.4.1.2 Supercell calculations

These were performed on single, zwitterionic and effectively isolated molecule of α-glycine, for comparison with the energy per molecule in the solid state, and thus allowing the deduction of the total intermolecular interaction energies in the crystal structure. Zero interaction between the nearest neighbouring cells was obtained by increasing the cell size and observing the change in the total energy. A cell size of 16 x 16 x 16 Å was found to break all intermolecular interactions, without giving rise to overly long computational times. Apart from a reduction in the number of k-points (to one, the gamma point), all other specifications for these calculations were the same as those applied for the corresponding crystal structure optimisation. The energy of the supercell (no atom relaxation) was compared to the solid-state energy of the optimised structure to obtain the lattice energy. In addition to this, an atom-only optimisation calculation was performed on a second supercell to obtain the neutral pseudo-isolated molecule; this energy was then compared to the energy of the optimised crystal structure (per molecule) in order to obtain an estimate of the sublimation energy. Note we did not attempt to obtain zero-point energy correction terms for these calculations, as this would require the computation of all phonon modes integrated over all k-space. Such a calculation is extremely expensive to perform and, as the correction term is expected to be very small, it was omitted in this work.

4.4.1.3 Dimer models

These models were constructed to estimate the individual contributions of each of the intermolecular interactions towards the lattice energy. All calculations performed were single-point energies (to avoid conversion to the more stable neutral gas-phase conformation) using the same basis-set cut-off and convergence criteria as reported for the full crystal structure optimisations; the k-point sampling grid was reduced to the gamma point. For dimer models 2 and 3 periodicity could be removed from the models by stretching the simulation cells along the 'a' and 'c' vectors, respectively, until the interaction beyond the immediate pair of molecules (i.e. a dimer) was destroyed. By comparing the energies thus obtained with the fully periodic model, an estimate of the σ-bond cooperativity effect can be obtained.

All the calculations carried out during this computational study were rigorously tested to ensure that the basis sets had reached an acceptable level of convergence. Consequently, the comparison of energies between the supercell and the rest of the calculations (*i.e.,* crystal structure and dimer calculations) is legitimate: we estimate any error incurred due to inconsistencies in basis sets to be within 0.01 – 0.003 eV (i.e., 0.2 – 0.07 kcal mol^{-1} or 0.8 – 0.3 kJ mol^{-1}) for the sublimation, lattice, proton transfer and hydrogen bond energies.

4.4.1.4 Isolated molecule calculations

In order to verify the results obtained from the PW-DFT dimer-model calculations, the same SPE calculations were also performed using localised basis sets as formalised in the Gaussian 98 code. We report hybrid density functional theory (DFT) calculations using the B3LYP functional. This set of calculations was performed using the 6-311G* basis set and all binding energies (quoted per bond) were corrected for basis set superposition error using the counterpoise (CP) correction proposed by Boys and Bernardi (1970).

4.4.2 Classical mechanical simulations

4.4.2.1 Ewald summations

To obtain the electrostatic contributions to the lattice energy (Ewald sum) the simulation package, GULP (Gale *et al.*, 2003) was used, taking the PW-DFT optimised structure and Mulliken charges as input.

4.4.2.2 Coulombic interaction energy calculations

These were done *via* an in-house program V_COULOMBIC, written to calculate the atom-atom pair electrostatic energy (Coulombic potential) contributions from a given set of coordinates and point charges.

4.5 Conclusions

The first aim of this study was to obtain a good understanding of the crystal structure of α-glycine, in terms of hydrogen bonding and thermodynamics, in the solid state, where molecules are in the zwitterionic form. The calculated sublimation and lattice energies themselves can give information about the overall energy of the hydrogen bonding present in the crystal structure, but we took this study a step further and also calculated the individual energies for the three interactions using different quantum mechanical and classical methods, which were found to be broadly consistent. Our calculated results for this test system therefore give confidence to the notion that amino acid systems can be reliably modelled by quantum mechanical and classical methods, which is crucial given the large variation in thermodynamic data recorded by different experimental methods (Table 4.1). It is well known that the intermolecular interactions present in the structures of biological molecules, such as amino acids, are in general classified as strong or moderate (e.g. salt bridge $N^+ – H\ldots {}^-O$) to weak hydrogen bonds (Desiraju *et al.*, 1999). We would therefore expect the values for their energies to fall in the 'textbook' range of *ca.* 4 – 15 kcal mol^{-1} (17 – 63 kJ mol^{-1}) (Jeffrey, 1997). It was therefore surprising and extremely interesting to conclude that the interaction energies for the α-glycine

crystal structure fall in a much wider range of energies of 1-30 kcal mol^{-1} (4-125 kJ mol^{-1}).

The second aim of the work presented in this chapter was to test different *ab initio* and classical methods in the study of zwitterionic amino acid crystal structures since there is very little available in the literature on this type of system. In summary, PW-DFT (CASTEP) calculated reliable geometries, lattice, sublimation and proton transfer energies, as well as Mulliken charges and individual interaction energies. The downside of this method is that calculations are generally slow and time-consuming. Consequently, during the course of this work a new procedure was developed to calculate the energies of the hydrogen bond interactions between amino acid zwitterion molecules with no appreciable loss of accuracy and considerable speed-up in calculation time. The first principles calculation provides an optimised geometry and set of Mulliken atomic charges. This information can then be utilised in a classical simulation (Ewald sum, GULP) to verify that the bulk of the lattice energy can be attributed to electrostatic intermolecular interactions; the individual atom-atom pair electrostatic interactions can then be further interrogated by a simple application of Coulomb's Law (V_Coulombic).

Finally, a clear correlation between the compressibility of the α-glycine intermolecular interactions, their geometries and their energies could be found, so that the strongest interaction in terms of geometry is also the least compressible, whereas the weakest interaction in terms of geometry presents the largest changes upon pressure. Therefore, these results show how the compressibility of the α-glycine crystal structure is not only dictated by kinetics but also by thermodynamics.

4.6 References

Allan, D. R.; Clark, S. J. (1999). *Phys. Rev*. **B60**, 6328-6334.

Berstein, J.; Davis, R. E.; Shimoni, L.; Chang, N. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555-1573.

Bisker-Leib, V.; Doherty, M. F. (2003). *Crystal Growth & Design* **3**, 221-237.

Boek, E. S.; Feil, D.; Briels, W. J.; Bennema, P. (1991). *J. Cryst. Growth*, **114**, 389-410.

Boys, S. F. & Bernardi, F, (1970). *Mol. Phys.*, **19**, 553.

CASTEP, Academic version 4.2, licensed under the UKCP-MSI Agreement, 1999; Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. (1992). *Rev. Mod. Phys.*, **64**, 1045.

Chickos, J. S. ; Acree, W. E. Jr. (2002). *J. Phys. Chem. Ref. Data*, **31**, 537.

Chisholm, J. A. ; Motherwell, S. ; Tulip, P. R. ; Parsons, S. ; Clark, S. J. (2005). *Crystal Growth & Design*, **5**, 1438-1442.

Császár, A. G. & Perczel, A. (1999). *Progr. .Biophys. Mol. Bio.* **71**, 243-309.

Dawson, A.; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L. (2005). *Crystal Growth & Design*. **5**, 1415-1427.

Day, G. M.; Price, S. L.; Leslie, M. (2003). *J. Phys. Chem.* **B107**, 10919-10933.

Derissen, J. L. ; Smlt, P. H. ; Voogd, J. (1977). *J. Phys. Chem.*, **81(15)**, 1474-1476.

Desiraju, G. R. ; Steiner, T. (1999). *The Weak Hydrogen Bond In Structural Chemistry and Biology*. Oxford University Press, New York.

Drebushchak, T. N.; Boldyreva, E. V.; Seryotkin, Yu. V.; Shutova, E. S. (2002). *J. Struct. Chem.* **43**, 835-842.

Frish, M. J. ; Trucks, G. W. ; Schlegel, H. B. ; Scuseria, G. E. ; Robb, M. A. ; Cheeseman, J. R. ; Zakrzewski, V. G. ; Montgomery, J. A. ; Stratmann, R. E. Jr ; Burant, J. C. ; Dapprich, S. ; Millam, J. M. ; Daniels, A. D. ; Kudin, K. N. ; Strain, M. C. ; Farkas, O. ; Tomasi, J. ; Barone, V. ; Cossi, M. ; Cammi, R. ; Mennucci, B. ; Pomelli, C. ; Adamo, C. ; Clifford, S. ; Ochterski, J. ; Peterson, G.A. ; Ayala, P. Y. ; Cui, Q. ; Morokuma, K. ; Malick, D. K. ; Rabuck, A. D. ; Raghavachari, K. ; Foresman, J. B. ; Cioslowski, J. ; Ortiz, J. V. ; Baboul, A. G. ; Stefanov, B. B. ; Liu, G. ; Liashenko, A. ; Piskorz, P. ; Komaromi, I. ; Gomperts, R. ; Martin, R. L. ; Fox, D. J . ; Keith, T. ; Al-Laham, M. A. ; Peng, C. Y. ; Nanayakkara, A. ; Gonzalez, C. ; Challacombe, M. ; Gill, P. M. W. ; Johnson, B. G. ; Chen, W. ; Wong, M. W. ; Andres, J. L. ; Head-Gordon, M. ; Replogle, E. S. ; Pople, J. A. *GAUSSIAN98*, revision A.9 ; Gaussian, Inc. : Pittsburgh, PA, 1998.

Gaffney, J. S. ; Pierce, R. C. ; Friedman, L. (1977). *J. Am. Chem. Soc.* **99**, 4293-4298.

Gale, J. D. & Rohl, A. L. (2003). *Molecular Simulation*, **29**, 291-341.

Jeffrey, G. A. (1997). *An Introduction to Hydrogen Bonding*, Oxford University Press, New York.

Jönsson, P. G.; Kvick, A. (1972). *Acta Cryst.* **B28**, 1827-33.

Kwon, O. Y.; Kim, S. Y.; No, K. T. (1995). *Bull. Kor. Chem. Soc.* **16**, 410-16.

Kwon, O. Y.; Kim, S. Y.; No, K. T.; Kang, Y. K.; Jhon, M. S.; Scheraga, H. A. (1996). *J. Phys. Chem.* **100**, 17670-17677.

Marsh, R. E. (1958). *Acta Cryst.* **11**, 654-63.

Materials Studio, 2001-2005, Accelerys Software Inc.

Merrill, L. & Bassett, W. A. (1974). *Rev. Sci. Instrum.* **45**, 290-294.

Moggach, S. A; Allan, D. R; Morrison, C. A; Parsons, S.; Sawyer, L. (2005). *Acta Cryst.* **B61**, 58-68.

Momany, F. A. ; Carruthers, L. M. ; Scheraga, H. A. (1974). *J. Phys. Chem.* **78**, 1621-1630.

Monkhorst, H. J. ; Pack, J. D. (1977). *Phys. Rev.* **B13**, 5188-5192.

Mulliken, R. S. J. (1955). *Chem. Phys.* **23**, 1833.

No, K. T.; Grant, J. A.; Scheraga, H. A. (1990). *J. Phys. Chem.* **94**, 4732-4739.

No, K. T.; Cho, K. H.; Kwon, O. Y.; Jhon, M. S.; Scheraga, H. A. (1994). *J. Phys. Chem.* **98**, 10742-9.

Perdew, J. P.; Burke, K.; Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865.

Popelier, P. L A; Joubert, L. (2002). *J. Am. Chem. Soc.* **124**, 8725-9.

Power, L. F.; Turner, K. E.; Moore, F. H. (1976). *Acta Cryst.* **B32**, 11-16.

Sanchez-Portal, D. ; Artacho, E. ; Soler, J. M. (1995). *Solid State Commun.* **95**, 685.

Segall, M. D. ; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. (2002). *J. Phys.: Condens. Matter*, **14**, 2717.

Shimura, K. (1951). *Nippon Nogei Kagaku Kaishi*, **24**, 412-16.

Svec, H. J. ; Clyde, D. D. (1965). *J. Chem. Eng. Data*, **10**, 151-152.

Takagi, S.; Chihara, H.; Seki, S. (1959). *Bull. Chem. Soc. Jap.* **32**, 84-8.

Tse, Y. C.; Newton, M. D.; Vishveshwara, S.; Pople, J. A. (1978). *J. Am. Chem. Soc.* **100**, 4329-31.

Volkov, A. & Coppens, P. (2004). *J. Comp. Chem.* **25**, 921-934.

Voogd, J. ; Derissen, J. L. ; van Duijneveldt. (1981). *J. Am. Chem. Soc.* **103**, 7701-7706.

Yu, D. ; Rauk, A. ; Armstrong, A. (1995). *J. Am. Chem. Soc.* **117**, 1789-1796.

Zhang, K. ; Chung-Phillips, A. (1999). *J. Chem. Inf. Comput. Sci.* **39**, 382-395.

Walker, M.; Morrison, C. A.; Allan, D. R. (2005). *Phys. Rev. B*, **77**, 224106/1-224106/9.

Chapter 5

Structural and Computational Study of L-α-aspartic Acid

5.1 Introduction

Aspartic acid, also known as L-aspartate (pH>5), is one of the two acidic naturally occurring non-essential amino acids which play an important role as general acids in enzyme active centres, as well as maintaining the solubility and ionic character of proteins. All of these functions are due to their polar side chain. The only other naturally occurring acidic amino acid is glutamic acid. The acidic and polar character arises from the carboxylic acid side group that both amino acids carry. Ionisable side groups can form metal and substrate binding and act as general acid-base catalysts. For example, three aspartic acids (D70, D65, D102 for E. Coli) and one glutamic acid (E20) are essential metal binding residues for pyrophosphate hydrolysis at the inorganic pyrophosphatase active site. Aspartic acid moves the coenzyme nicotinamide adenine dinucleotide, NADH, molecules from the main body of the cell to its mitocondria, where it is used to generate adenosine triphosphate, ATP, the fuel that powers all cellular activity. In addition to this, aspartic acid removes excess toxins from the cells, particularly ammonia, which is very damaging to the brain and nervous system as well as the liver. Aspartic acid is widely present in proteins and acts both as a base and an acid in the citrate synthase reaction of the citric acid cycle, also known as Krebs Cycle, which is the means of energy production for all animal species. Aspartic acid, which was first isolated in 1868 from the legumin in plant seeds, can be found especially in young sugar cane and sugar-beet molasses and it is used as a treatment for chronic fatigue.

The crystal structure of L-α-aspartic acid, obtained by X-ray diffraction techniques, was initially reported by Derissen *et al.* (1968). L-α-aspartic acid crystallises in the monoclinic $P2_1$ space group and has a total of two molecules in the unit cell with one in the asymmetric unit. The cell parameters reported in this investigation were a = 15.1, b = 6.9, c = 5.1 Å, and β = 96°. Derissen *et al.* described the crystal structure of L-α-aspartic acid in terms of layers of molecules: "*we can formally distinguish two 'layers' of L-aspartic acid molecules; one molecular 'layer' is transformed into the second by a screw axis. The two 'layers' are connected by hydrogen bonds (between charged atoms) in which the N atom is involved. In one 'layer' the molecules are linked together in zig-zag chains, and the chains are also*

connected via the N-hydrogen bonds" (Derissen *et al.*, 1968). In order to determine the effect of compression on the intermolecular interactions and molecular packing, the crystal structure of L-α-aspartic acid was determined at a number of pressures, in the range from 0 to 5.8 GPa, using single-crystal X-ray diffraction with a Merrill-Bassett diamond-anvil cell. So that the observed structural changes could be described in a more detailed manner, the description by Derissen *et al.* (1968) of the L-α-aspartic acid crystal structure will be expanded upon using graph set notation and Hirshfeld surfaces to denote and visualise both the connectivity and relative strengths of the intermolecular interactions.

In addition to the experimental work, a computational study similar to the one presented for cyclopropylamine and glycine in chapters 3 and 4 was carried out on the crystal structure of L-α-aspartic acid, to determine the energetics of the hydrogen bonds present in the structure so that these could be related to the bond compressibilities.

5.2 Experimental study of L-α-aspartic acid

5.2.1 Crystal growth

L-α-aspartic acid was obtained from the Aldrich Chemical Company. Colourless crystals, in the form of needles, were grown from water by slow evaporation at ambient temperature. A small, block-shaped crystal was selected from the batch of crystals and loaded into a diamond anvil cell.

5.2.2 High pressure X-ray crystallography

The high-pressure experiments were performed using a Merrill-Bassett diamond anvil cell (Merrill & Bassett, 1974), which has a half-opening angle of 40°, and was equipped with 600 μm diamond culets and a tungsten gasket. A 250 μm hole was drilled through the tungsten gasket in order to accommodate the single crystal of L-α-

aspartic acid. A 4:1 mixture of methanol and ethanol was used as a hydrostatic pressure transmitting medium. A small ruby chip was also loaded into the cell, so that its fluorescence spectrum could be used to yield the sample pressure. Pressure measurement was carried out by excitation of the ruby R_1 and R_2 fluorescence line emission with a 632.417 nm line from a He-Ne laser and the resulting ruby fluorescence spectrum was detected with a Jobin-Yvon LabRam 300 Raman spectrometer.

Diffraction data were collected on a Bruker SMART APEX diffractometer with graphite-monochromated Mo-Kα radiation (λ = 0.71073 Å). Data collection (SMART; Bruker-AXS, 1991-2001) and processing procedures (GEMINI; Bruker-AXS, 1999) for the high-pressure experiments were as described by Dawson *et al.* (2004). Data collections were taken in steps from 0 GPa up to a final pressure of 5.8 GPa. Integrations were carried out using the program SAINT (Bruker-AXS, 2003) which resulted in a completeness range of 38.3 to 24.5%. Absorption correction was undertaken with the programs SORTAV (Blessing, 1987, 1989) and ABSORB (Angel, 2004) at all sample pressures. The unit cell dimensions ranged from a = 7.6000(17), b = 6.9634(17), c = 5.135(3) Å, and β = 99.76(4)° at 0 GPa and a = 7.364(3), b = 6.059(2), c = 5.002(4) Å, and β = 103.42(6)° at 5.8 GPa and the space group was found to remain monoclinic, $P2_1$. No discontinuities were observed in unit cell parameters or their first derivatives with pressure, other than what could be expected from a continuous smooth compression and consequently there was no evidence for a structural phase transition.

Refinements were carried out against $|F|^2$ using all data (CRYSTALS, Betteridge *et al.*, 2003). Due to the limited completeness of the high-pressure data sets, all 1,2 distances were restrained to the values observed in the ambient pressure structure, and all carbon, nitrogen and oxygen atoms were refined with isotropic displacement parameters. Additionally, restraints were applied to the thermal and vibrational motion of the atoms.

Pressure/ GPa	**0**	**1.4**	**2.9**	**3.8**	**5.8**
			Crystal data		
Chemical formula			$C_4H_7NO_4$		
M_r			133.10		
Cell setting, space group			Monoclinic, $P12_11$		
a, b, c (Å)	7.6000 (17), 6.9634 (17), 5.135 (3)	7.5297 (12), 6.7387 (13), 5.0619 (17)	7.462 (2), 6.503 (2), 5.041 (3)	7.4214 (9), 6.2863 (9), 5.0249 (14)	7.364 (3), 6.059 (2), 5.002 (4)
β (°)	99.76 (4)	101.96 (3)	102.84 (6)	103.34 (2)	103.42 (6)
V (Å^3)	267.82 (16)	251.27 (11)	238.50 (19)	228.10 (8)	217.1 (2)
Z			2		
D_x (Mg m^{-3})	1.650	1.759	1.853	1.938	2.036
Radiation type			Mo $K\alpha$		
No. of reflections for cell parameters	221	198	213	213	267
θ range (°)	3–23	3–20	3–23	3–22	3–23
μ (mm^{-1})	0.15	0.16	0.17	0.18	0.18
Temperature (K)	293	293	293	293	293
Crystal form, colour			Block, colourless		
			Data collection		
Diffractometer			Bruker SMART		
Scan Method			ω scans		

Pressure/ GPa *(Cont.)*	**0**	**1.4**	**2.9**	**3.8**	**5.8**
Absorption correction			Multi-scan (based on symmetry-related measurements)		
$T_{min,}$ T_{max}	0.587, 1.000	0.690, 1.000	0.696, 1.000	0.725, 1.000	0.492, 1.000
No. of measured, independent and observed parameters	164, 161, 97	148, 146, 113	143, 141, 109	141, 138, 108	132, 129, 76
Criterion for observed reflections			$I > 2.00u(I)$		
R_{int}	0.145	0.092	0.094	0.101	0.154
θ_{max} (°)	23.3	23.1	23.3	23.2	23.2
Range of h, k, l	$-8 \rightarrow h \rightarrow 8$	$-8 \rightarrow h \rightarrow 8$	$-8 \rightarrow h \rightarrow 8$	$-8 \rightarrow h \rightarrow 8$	$-8 \rightarrow h \rightarrow 8$
	$-7 \rightarrow k \rightarrow 7$	$-7 \rightarrow k \rightarrow 7$	$-7 \rightarrow k \rightarrow 7$	$-6 \rightarrow k \rightarrow 6$	$-6 \rightarrow k \rightarrow 6$
	$-2 \rightarrow l \rightarrow 1$	$-2 \rightarrow l \rightarrow 1$	$-2 \rightarrow l \rightarrow 1$	$-2 \rightarrow l \rightarrow 1$	$-1 \rightarrow l \rightarrow 2$
			Refinement		
Refinement on			F^2		
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.087, 0.192, 0.95	0.078, 0.187, 1.04	0.073, 0.158, 1.04	0.073, 0.171, 0.96	0.120, 0.265, 0.99
No. reflections, No. parameters	142, 37	146, 40	141, 40	138, 40	112, 40
H-atom treatment			Mixture of independent and constrained refinement		
Weighting scheme*: P(1), P(2)	0.0206, 2.93	0.0563, 1.62	0.00, 1.66	0.00, 2.32	0.00, 6.16
$(\Delta/\sigma)_{max}$			<0.0001		
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.41, –0.35	0.31, –0.30	0.42, –0.39	0.38, –0.31	0.67, –0.61

Computer programs: *SMART* (Siemens, 1993); *SAINT* (Siemens, 1995); *CRYSTALS* (Betteridge *et al.*, 2003);. * where $w = 1/[\sigma^2(F^2) + (P(1)p)^2 + P(2)p$.

Table 5.1: Crystallographic data for L-α-aspartic acid at increasing pressures.

The L-α-aspartic acid coordinates of Derissen *et al.* (1968) were refined against these data to yield a conventional *R*-factor of 0.087 for 142 data with $I > 2\sigma(I)$ when data was collected inside the diamond anvil cell at 0 GPa. The *R*-factor values obtained for the refinements at the different pressures, from 0 to 5.8 GPa, fell in the range from 0.073 to 0.12. The aim of this zero-pressure experiment was simply to study the effect of the diamond anvil cell on the quality of the diffracted intensities and the subsequent structural refinement. The resulting crystallographic data are shown in Table 5.1. It was found that the refined structure obtained with the DAC at 0 GPa is consistent (within standard error) with those of Derissen *et al.* (1968) reported at ambient conditions (i.e. mounted on a fibre). However, the position of the hydrogen atoms differ as they were placed geometrically during the refinement procedure and their positions were not refined due to the limited quality of data, with respect to counting statistics and overall completeness. Nevertheless, the H atoms were allowed to move (ride) with the heavy atoms, contributing in this way to the refinement of the model. Consequently, the difficulty of locating the hydrogen atoms during the X-ray diffraction analyses led to the identification of hydrogen bonds from the D…A distances (D = donor, A = acceptor) alone and only the distances between the nitrogen and oxygen atoms involved in the interactions will be discussed. The values for these distances at 0 GPa are in good agreement with those reported by Derissen *et al.* (1968) at room temperature.

5.2.3 Results and discussion

5.2.3.1 Analysis of the crystal structure of L-α-aspartic acid at ambient pressure

The numbering scheme reported by Derissen *et al.* (1968) in the Cambridge Structural Database (refcode LASPRT) will be used to describe the structure. The molecule of L-α-aspartic acid is in the zwitterionic form, which is the most stable form for amino acids in the solid state, whereas the neutral form is preferred in the gas phase. Only one of the two carboxylic groups in the L-α-aspartic acid molecule is ionised, whereas the second carboxylic group is in the neutral form. There is one molecule of L-α-aspartic acid in the asymmetric unit and there are two in the unit cell,

which are bridged by short intermolecular hydrogen bonds forming subunits, which are invariant parts of the structure. The subunit molecules are in a head-to-head arrangement, whereas molecules of different subunits are in a tail-to-tail conformation. The OH groups present in the subunits are arranged *trans* with respect to the central part of the subunit, where the interaction between the two molecules take place (see Figure 5.1), with angles of 173.1°. The combination of the subunits builds the crystal structure of L-α-aspartic acid [Figure 5.1(a)-(b)], which can be described as layers of molecules arranged parallel to the (1 0 1) planes and stacked along the c axis. The formation of subunits and layer-like arrangements are common features in the monoclinic crystal structures of many amino acids and peptides. In glycine (GLYCIN; Iitaka, 1960) for example, subunits similar to those described for aspartic acid, form layers parallel to the *ac*-face diagonal of the α-phase [Figure 5.2(a)]. The subunits in the β phase of glycine (GLYCIN02; Marsh, 1958) form layers parallel to the c axis and are stacked along the a axis [Figure 5.2(b)]. In α-glycylglycine (GLYGLY; Biswas *et al.*, 1968) the molecular subunits form corrugated layers which stack parallel to the *a* crystallographic direction [Figure 5.2(c)]. In all these cases, the layers are normally formed parallel to the plane which contains the longest cell dimension and are stacked parallel to the shortest cell dimension, which in most cases has a value of around 5.1 Å.

Each molecule of L-α-aspartic acid participates in the formation of four hydrogen bonds (Table 5.2). Thus the NH_3^+ group actively forms three N-H…O hydrogen bonds, whereas the COO^- group forms a O-H…O hydrogen bond with the COOH group [Figure 5.3(a)-(e)]. One of the unit cell dimensions exhibits a value of around 5.1 Å, which is due to a head-to-head chain motif formed by the N1-H4...$O2^i$ interaction (2.862 Å, 175°). This is a common feature of the crystal structures of amino acids, in which the motif can be different from the head-to-head type (L-serine; Moggach *et al.*, 2005). This N1-H4…$O2^i$ hydrogen bond connects molecules of L-α-aspartic acid to form chains of molecules parallel to the c-axis [Figure 5.3(a)], described as *C(5)* in graph set notation (Etter *et al.*, 1990; Bernstein *et al.*, 1995).

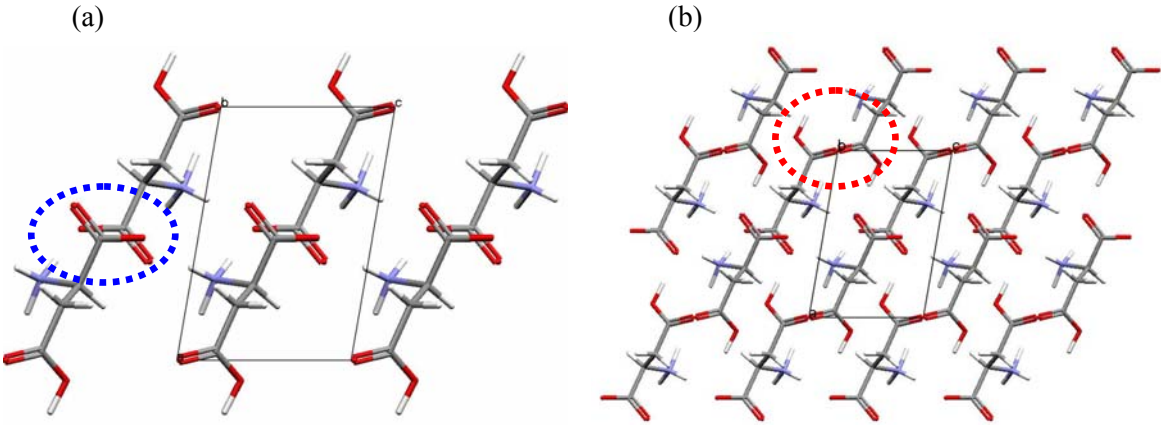


Figure 5.1: View of the layers formed by subunits linked *via* (a) head-to-head (intra-subunits) (●●●), and (b) head-to-head (inter-subunits) and tail-to-tail (inter-subunits) interactions (●●●), along the c-axis in the crystal structure of L-α-aspartic acid at ambient pressure (LASPRT, Derissen *et al.*, 1968).

Pressure/GPa	**Derissen *et al.* (1968)**	**0**	**0.2**	**1.4**	**2.9**	**3.8**	**5.8**
D(N1…O2)[i]/ Å	2.862	2.90(4)	2.85(7)	2.78(3)	2.78(3)	2.77(3)	2.76(6)
D(N1…O2)[ii]/ Å	2.800	2.82(3)	2.82(5)	2.79(2)	2.75(2)	2.71(2)	2.65(4)
D(N1…O3)[iii]/ Å	2.813	2.78 (2)	2.74(4)	2.81(2)	2.830(18)	2.806(19)	2.68(3)
D(O4…O1)[iv]/ Å	2.577	2.55(2)	2.52(4)	2.53(2)	2.495(18)	2.477(17)	2.47(3)
D(N1…O1)/ Å	2.944	2.93(3)	2.84(5)	2.80(3)	2.713(18)	2.70(3)	2.56(5)
D(O4…O3)/ Å	3.130	3.07(3)	2.92(6)	3.06(3)	3.06(3)	3.02(3)	2.90(4)

i) *x, y, 1+z*; ii) *1-x,1/2+ y ,1-z*; iii) *-x, 1/2+y, 2-z*; iv) *-1+x, -y, z*.

Table 5.2: Hydrogen bonding parameters in L-α-aspartic acid at increasing pressures.

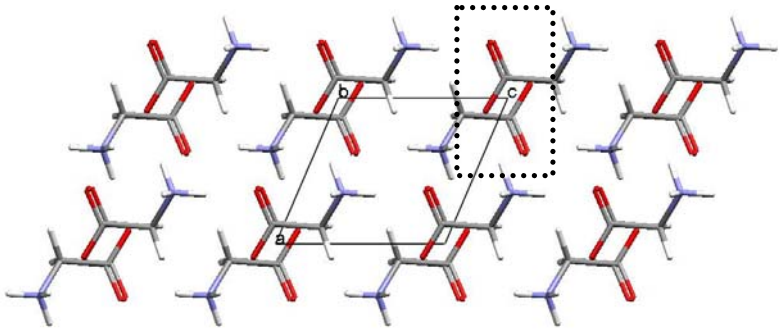
The N1-H6…O3 interaction (2.813 Å, 173°) links L-α-aspartic acid molecules to form *C(6)* chains, which are generated by the 2_1 screw axis parallel to the b-axis [Figure 5.3(b)]. The torsion angle formed by two different consecutive hydroxyl groups within the *C(6)* chain is 175.2°. This means that the hydroxyl groups present in the different molecules adopt a *trans* arrangement with respect to other hydroxyl groups along the 2_1 screw axis. The combination of the *C(5)* and *C(6)* chains, formed

by the N1-H4…O2^{i} and N1-H6…O3 interactions, form $R_3^4(20)$ ring motifs, which stack along the crystallographic c-axis [Figure 5.4(a)].

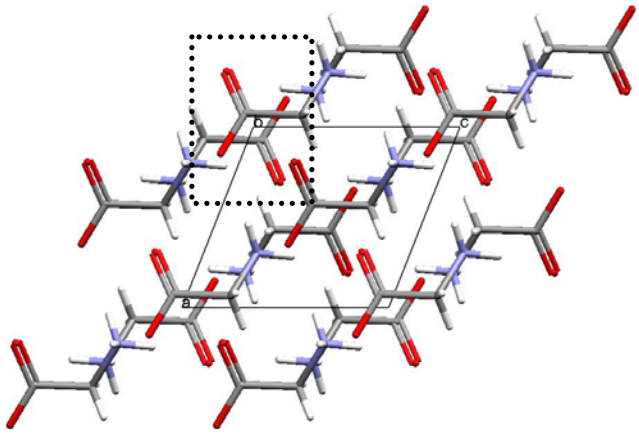
The third interaction in which the NH_3^+ group is involved is the N1-H5…O2^{ii} hydrogen bond (2.800 Å, 179°). This interaction holds molecules together to form the subunit 1 [see Figure 5.1(a)] and forms *C(5)* chains parallel to the b-axis [Figure 5.3(c)]. Additionally, the hydrogen bond N1-H5…O2^{ii}, together with N1-H4…O2^{i}, forms $R_3^3(14)$ rings [Figure 5.4(b)], which link molecules together to form corrugated layers parallel to the *bc* face. The combination of the N1-H5…O2^{ii} and N1-H6…O3 interactions give rise to $R_3^4(20)$ ring motifs [Figure 5.4(c)], which also contribute to the formation of corrugated layers along the b-direction. Finally, the three hydrogen bonds formed by the NH_3^+ group form $R_3^5(22)$ ring motifs, which connect molecules to form subunits but also inter-link subunits [Figure 5.4(d)].

The remaining hydrogen, O4-H1…O1, interaction (2.577 Å, 175°), which forms *C(7)* chains that run parallel to the a-axis, is created by the OH group [Figure 5.3(d)]. This interaction is involved in the formation of $R_3^4(19)$ and $R_3^4(15)$ [Figure 5.4(e)] ring motifs, with the other three hydrogen bonds formed by the NH_3^+ group. The O4-H1…O1 hydrogen bond is also involved in both connecting pairs of molecules to form subunits and the linking of the subunits within the hydrogen bond network of the crystal structure.

(a)



(b)



(c)

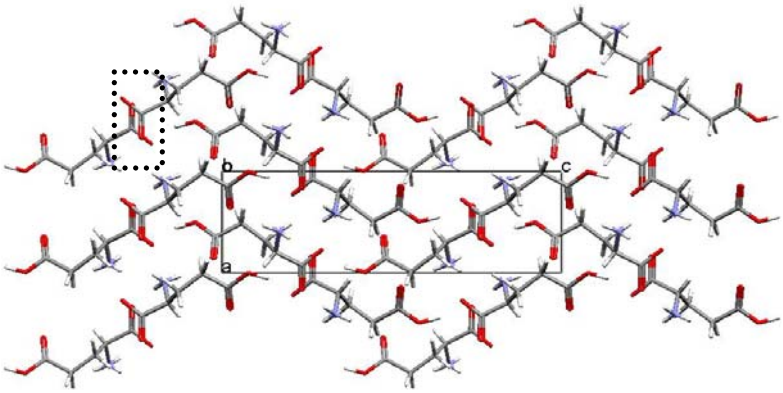


Figure 5.2: Views of the layers formed by the different subunits formed in the crystal structure of (a) α-glycine (Iitaka, 1960), (b) β-glycine (Marsh, 1958), and (c) α-glycylglycine (Biswas *et al.*, 1968). The dotted boxes show the bonds linking molecules within the subunits.

Apart from the four strong hydrogen bonds, there is a fourth N1-H4…O1 interaction (2.944 Å, 87°) and a second O4…O3 (3.130 Å, 80°) interaction, which are very weak. The N1-H4…O1 hydrogen bond connects molecules between subunits, forming C(4) chains of molecules parallel to the b-axis, with the OH groups of the molecules pointing in antiparallel directions [Figure 5.4(f)]. The O4…O3 interaction inter-links subunits to form the C(6) chains [Figure 5.4(g)]. The N1-H4…O1 interaction forms $R_2^2(8)$ and $R_2^3(10)$ ring motifs with N1-H6…O3 and O4-H1…O1 hydrogen bonds, which are positioned at the corners of the unit cell, whereas the O4…O3 interaction, along with the N1-H6…O3, O4-H1…O1 and N1-H4…O1 hydrogen bonds, forms $R_2^1(8)$ and $R_2^3(13)$ rings. The $R_2^2(8)$, $R_2^1(8)$ and $R_2^3(13)$ ring motifs are combined to form one of the corrugated molecular layers [Figure 5.4(h)-(j)].

In addition to the N-H…O and O-H…O interactions, there is a weak C3-H2…O4 hydrogen bond (3.318 Å, 136°), which links molecules to form *C(4)* chains parallel to the b-axis [Figure 5.4(j)]. This interaction, although weak, plays an important role in the structure as it interconnects neighbouring L-α-aspartic acid molecules. The importance of weak C-H…O weak interactions and their role in stabilising the secondary and tertiary structures of proteins is well documented (Desiraju *et al.*, 1999).

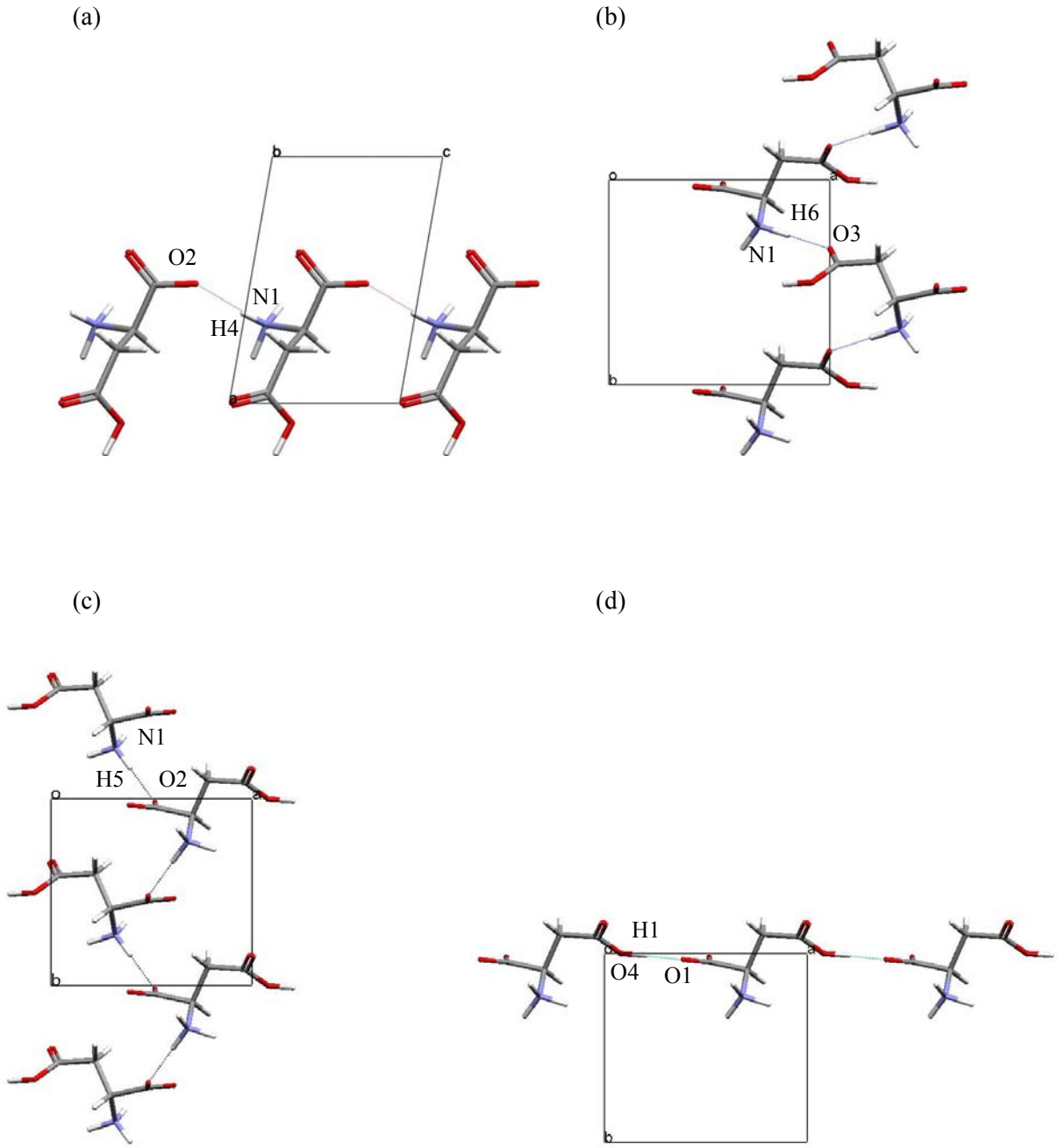
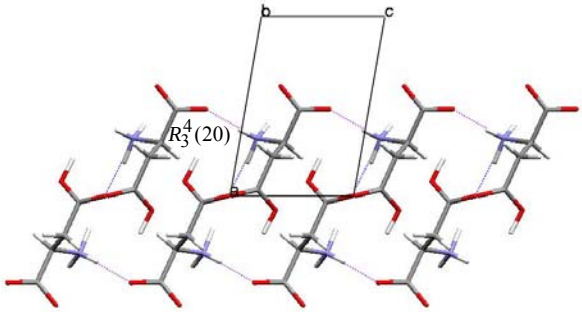
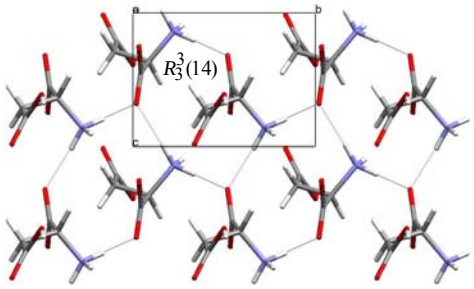


Figure 5.3: Views of the chains formed by the four hydrogen bonds present in the crystal structure of L-α-aspartic acid at ambient pressure (LASPRT, Derissen *et al.*, 1968): (a) N1-H4…O2^{i} (•••), (b) N1-H6…O3 (•••), (c) N1-H5…O2^{ii} (•••), (d) O4-H1…O1 (•••).

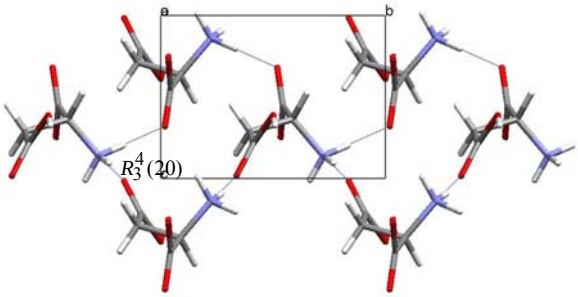
(a)



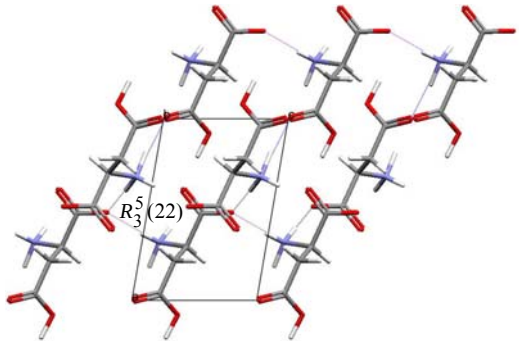
(b)



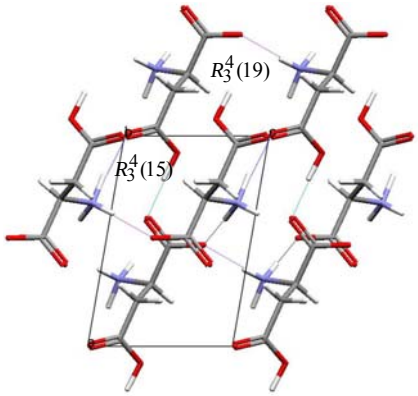
(c)



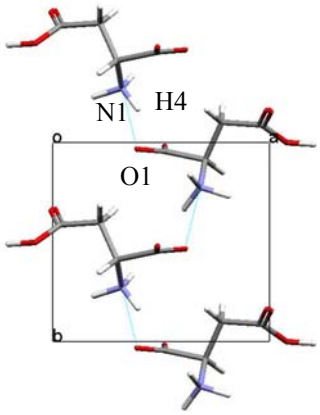
(d)



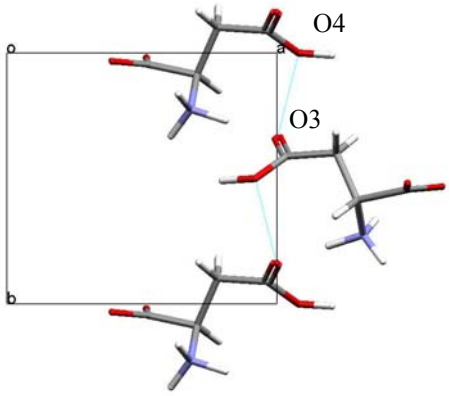
(e)



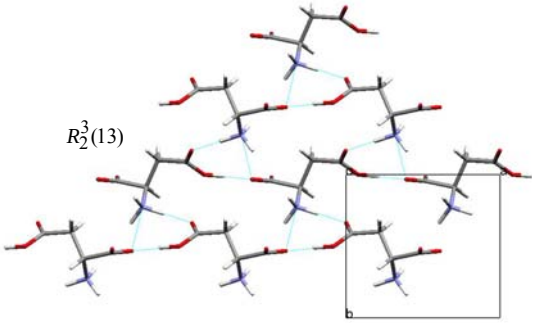
(f)



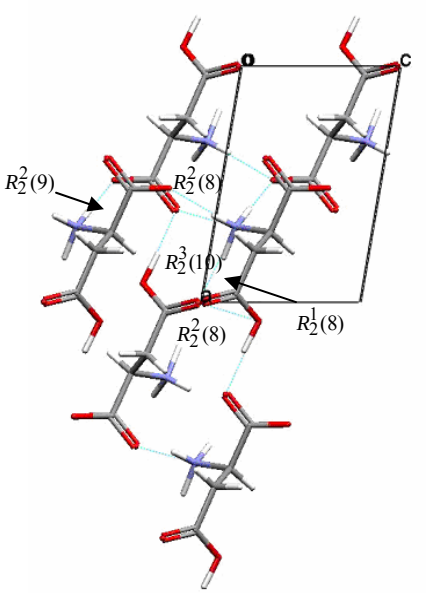
(g)



(h)



(i)



(j)

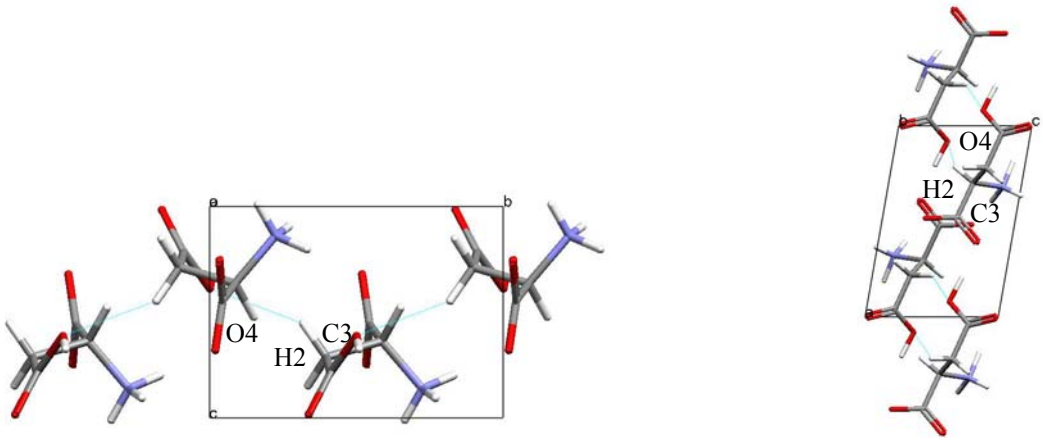


Figure 5.4: The crystal structure of L-α-aspartic acid at ambient pressure (LASPRT, Derissen *et al.*, 1968). (a) View of the layer of $R_3^4(20)$ ring motifs formed by the *C(5)* and *C(6)* chains linked by N1-H4…O2^{i} (●●●)and N1-H6…O3 (●●●) hydrogen bonds respectively, which run along the c-axis. (b) View of the layers of $R_3^3(14)$ rings formed by the N1-H4…O2^{i} and N1-H5…O2^{ii} (●●●) interactions, giving rise to corrugated layers, parallel to the b-axis. (c) View of the layer of $R_3^4(20)$ ring motifs, also present between corrugated layers, formed by N1-H5…O2^{ii} and N1-H6…O3 hydrogen bonds, along the b-axis. (d) Layer of $R_3^5(22)$ rings formed by the three hydrogen bonds involving NH_3^+, and which link subunits along the c-axis. (e) Formation of $R_3^4(19)$ and $R_3^4(15)$ ring motifs linked by the strong O4-H1…O1 (●●●) interaction together with the NH_3^+ hydrogen bonds. (f) Two *C(5)* chains linked parallel to the b-axis by N1-H4…O1. (g) O4…O3 interactions running along the b-axis. (h) View of $R_2^2(8)$ ring motifs formed by O4-H1…O1, N1-H6…O3 and N1-H4…O1 interactions, giving rise to a hydrogen bonded network of molecules, parallel to the *ab*-face. (i) Formation of $R_2^1(8)$, $R_2^2(8)$, $R_2^2(9)$, $R_2^3(10)$ and $R_2^3(13)$ ring motifs, which involve all the interactions present in the crystal structure. These rings connect subunits to form layers of molecules parallel to the *ac*-face. (j) *C(4)* chains linked by weak C3-H4…O4 hydrogen bonds run along the b-axis and connect molecules between subunits.

5.2.3.2 The effect of pressure on the intra-molecular bond distances and bond angles

On pressure increase from 0 to 5.8 GPa the intra-molecular bond distances and bond angles did not change significantly with respect to experimental uncertainties. However, some significant changes were observed in the torsional angles (Table 5.2, Figure 5.5). All the changes in the torsional angles were in the range of 15 to 19°, making the different parts of the molecule more or less planar: i.e. the C3-C4-O3-O4 group is clearly increasing its planarity (see Figure 5.6), whereas in the rest of the molecule the planarity is decreasing (e.g. C2-C3-C4-O3, N1-C2-C3-C4). When pressure is applied to the structure, the L-α-aspartic acid molecules increase their planarity and reorient to become more parallel within the molecular chains (Figure 5.6)

Pressure / GPa	**∠(O2-C1-C2-C3)/°**	**∠(C2-C3-C4-O3)/°**	**∠(N1-C2-C3-C4)/°**	**∠(O2-C1-C2-N1)/°**
0	91(2)	131(2)	60(3)	149(2)
1.4	104(2)	123(2)	59(2)	140(2)
2.9	103(2)	119(2)	57(2)	137(2)
3.8	104(2)	114(2)	54(2)	136(2)
5.8	107(4)	111(4)	50(4)	130(4)

Table 5.2: Changes in selected torsional angles of the L-α-aspartic acid molecule at increasing pressures.

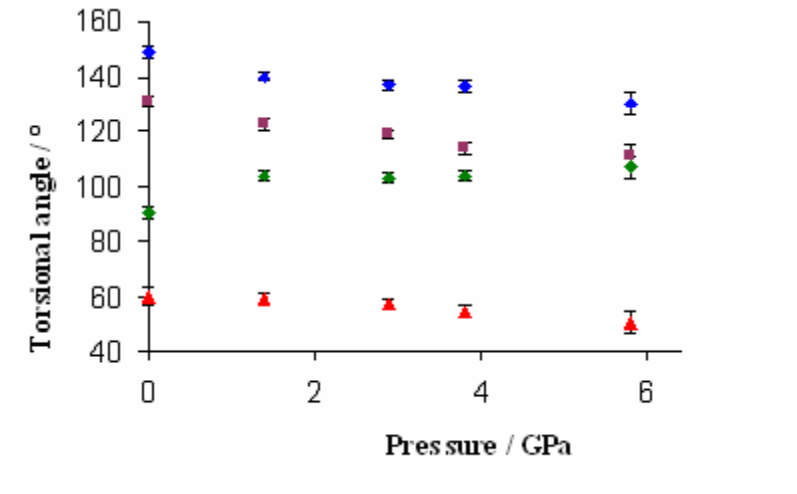


Figure 5.5: Intramolecular torsional angles (°) for the L-α-aspartic acid molecule, versus pressure (GPa) in the crystal structure of L-α-aspartic acid. [♦-(O2-C1-C2-C3), ■-(C2-C3-C4-O3), ▲-(N1-C2-C3-C4), ♦-(O2-C1-C2-N1)].

(a) (b)

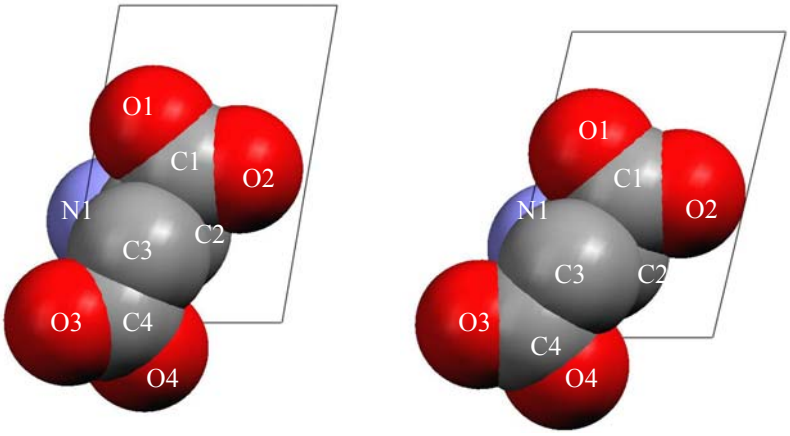


Figure 5.6: (a, b) Space-filling representations of the L-α-aspartic acid molecule in the crystal structure at 0 and 5.8 GPa. From the plots it can be seen how the C3-C4-O3-O4 group increases its planarity with increasing pressure.

5.2.3.3 The effect of pressure on hydrogen bonding

It was found that when pressure was applied to the L-α-aspartic acid crystal structure, all the intermolecular interactions changed significantly. From Table 5.3 it can be seen that all the D…A hydrogen bond distances decreased their value on pressure increase from 0 to 5.8 GPa. Thus, the N1…O2^{i} interaction decreased its distance from 2.82(3) to 2.64(4) Å (6.3%), whereas the N1…O3 and N1…O2^{ii} interactions decreased their distances from 2.78(2) and 2.90(4) Å to 2.68(3) and 2.76(6) Å, respectively (3.6 and 4.9%). Finally, the O4…O1 interaction decreased from 2.55(2) to 2.47(3) Å (3.1%), whereas the weak N1…O1 interaction shortened from 2.93(3) to 2.56(5) Å (12.7%) (Figure 5.7). From these results, as can be anticipated, there is a clear correlation between the initial hydrogen bond distance and the compressibility of the interaction, so that the interaction with the least compression is the shortest (i.e., the strongest) and the interaction with the largest compression is the longest (i.e., the weakest). In comparison, the minimum COOH…¯OOC distance found in the CSD is 2.460 Å (NOPSUG, Llamas-Saiz *et al.*, 1994) for a supramolecular complex, which is only some 0.01 Å shorter than the O4…O1 distance reported here.

Pressure/GPa	**Derissen *et al.* (1968)**	**0**	**1.4**	**2.9**	**3.8**	**5.8**
D(N1…O2)^{i}/ Å	2.862	2.90(4)	2.78(3)	2.78(3)	2.77(3)	2.76(6)
D(N1…O2)^{ii}/ Å	2.800	2.82(3)	2.79(2)	2.75(2)	2.71(2)	2.65(4)
D(N1…O3)^{iii}/ Å	2.813	2.78 (2)	2.81(2)	2.830(18)	2.806(19)	2.68(3)
D(O4…O1)^{iv}/ Å	2.577	2.55(2)	2.53(2)	2.495(18)	2.477(17)	2.47(3)
D(N1…O1)/ Å	2.944	2.93(3)	2.80(3)	2.713(18)	2.70(3)	2.56(5)
D(O4…O3)/ Å	3.130	3.07(3)	3.06(3)	3.06(3)	3.02(3)	2.90(4)

i) *x, y, 1+z*; ii) *1-x,1/2+ y ,1-z*; iii) *-x, 1/2+y, 2-z*; iv) *-1+x, -y, z*.

Table 5.3: Hydrogen bonding parameters in L-α-aspartic acid at increasing pressures.

Some of the N…O interactions decreased their distance to around 2.65 Å, which is the distance obtained in the high pressure structures of other amino acids, before a phase transition takes place. For example Moggach *et al.* (2005) reported that a N…O

distance decreased from 2.691(13) to 2.65(4) Å immediately prior to a phase transition. The hydrogen bond became shorter than the minimum N…O distance of 2.661 Å (DUSMAF; Suresh *et al.*, 1986) found in amino acids in the Cambridge Structural Database. In the compressibility study of L-α-aspartic acid it was found that the N1…O2^{ii} hydrogen bond decreases its length from 2.82(3) to 2.65(4) Å on pressure increase. In addition to this, and more strikingly, the N1…O1 interaction shortens from 2.93(3) to 2.56(5) Å, which is shorter than the minimum N…O distance found in the Cambridge Structural Database for similar compounds.

An obvious consequence of the decreased D…A distances is that the structure at 5.8 GPa is more dense and compact than that at 0 GPa. When pressure increased from 0 to 5.8 GPa most of the changes regarding the subunit formation took place not within subunits, but between them, giving a strong indication of their rigid nature. The subunits translated with pressure in order to minimise the space between them and to flatten the molecular planes [Figure 5.8(a)]. This is also illustrated in the form of space-filling plots [Figure 5.8(b)], which show how the voids between molecules decrease in volume on pressure increase. In addition to the closure of the voids, from Figure 5.9 it is possible to compare the conformations of the molecules at 0 and 5.8 GPa. The closure of the voids in the crystal structure can be investigated by studying the behaviour of the different hydrogen bonded rings on pressure increase. For example, the holes in the $R_3^4(20)$ and $R_3^3(14)$ ring motifs decrease their size considerably [Figure 5.10(a)-(d)], as a consequence of the contribution from the N1…O2^{i} interaction and its significant shortening. In general, all the rings decrease their size when going from 0 to 5.8 GPa but the most affected ring motifs are those formed by the interactions with the largest changes (i.e., N1…O2^{i} and N1…O2^{ii}).

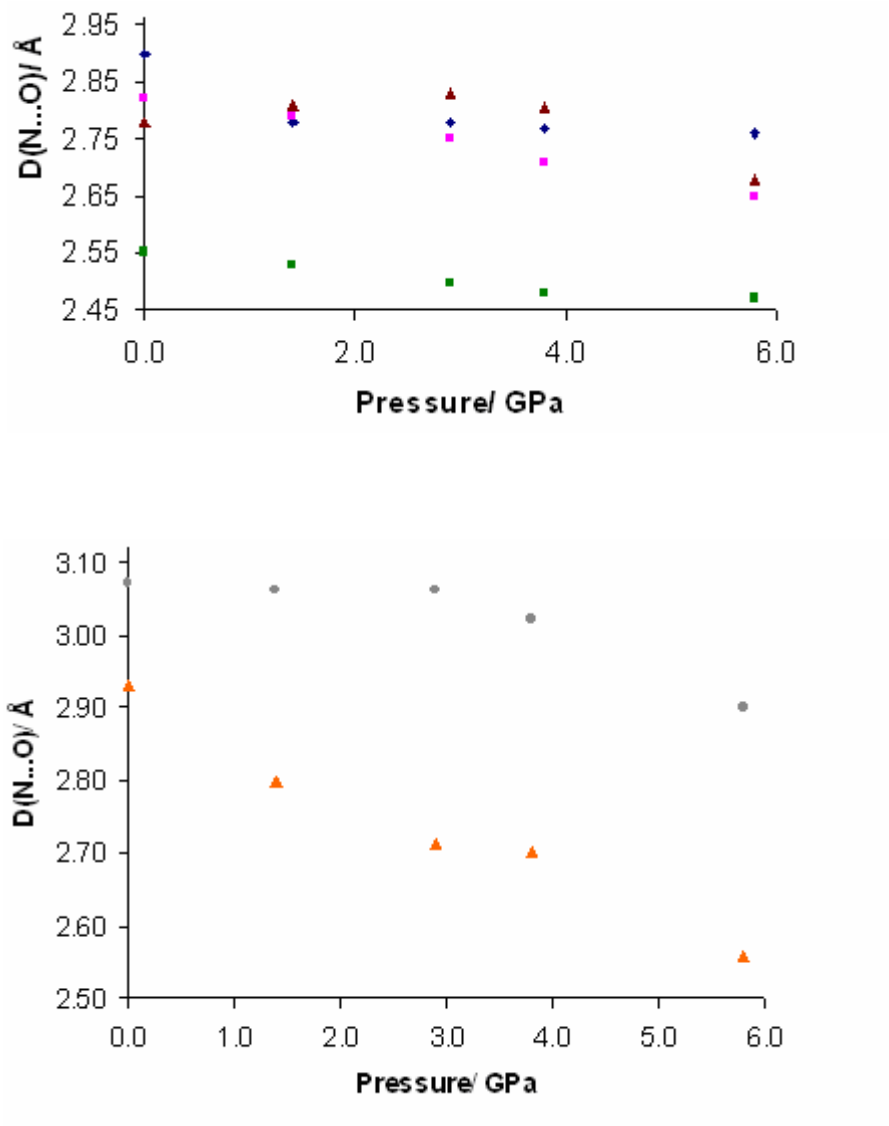


Figure 5.7: Intermolecular interactions distances (Å) for the different hydrogen bonds and close contacts formed by the NH_3^+ and OH groups, versus pressure (GPa) in the crystal structure of L-α-aspartic acid. [♦-(N1…O2)I, ■-(N1…O2)ii, ▲-(N1…O3)iii, ■-(O4…O1)iv, ●-(O4…O3), ▲-(N1…O1)].

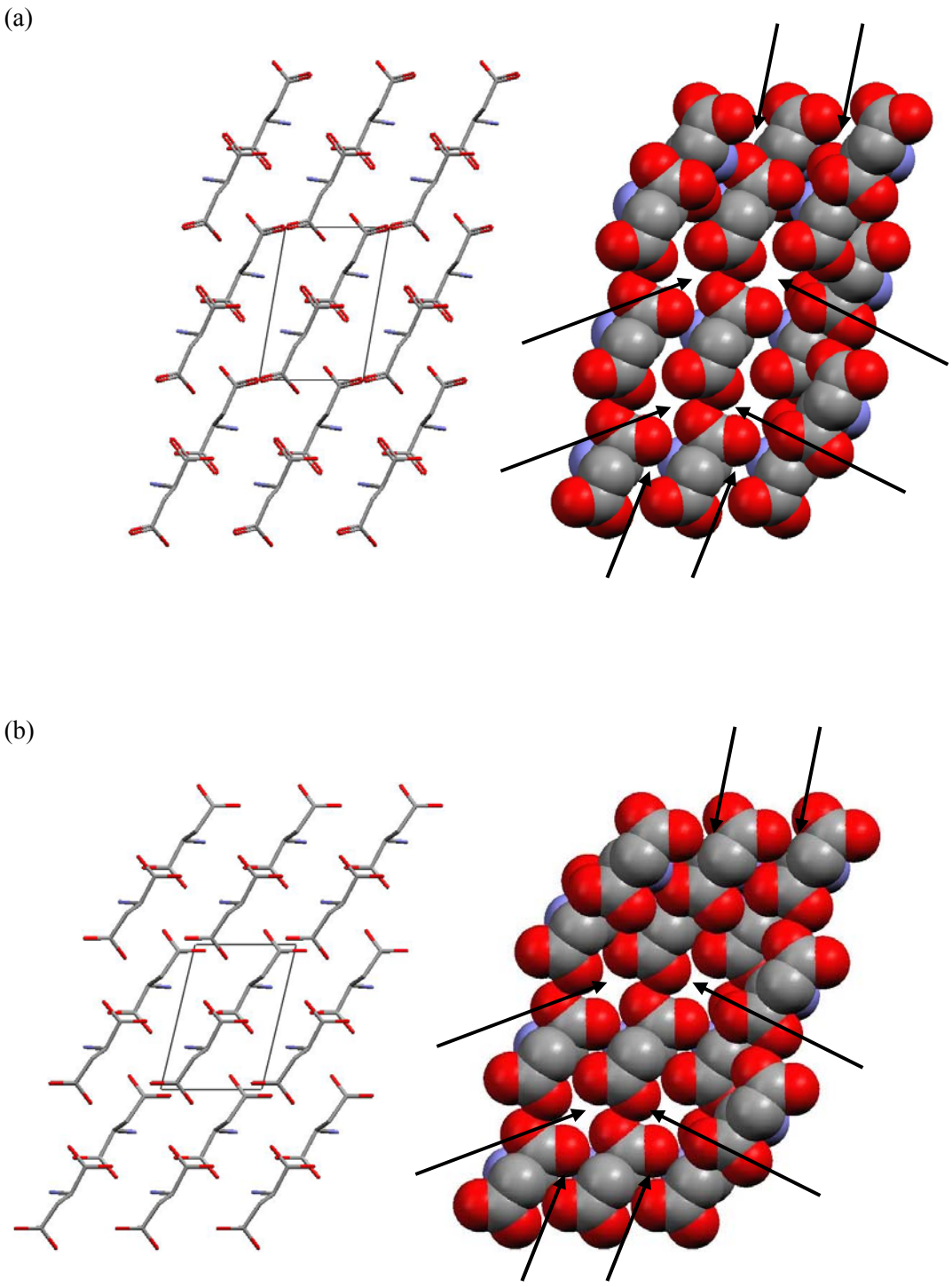


Figure 5.8: (a, b) Space-filling representations of the L-α-aspartic acid structure at 0 and 5.8 GPa. Only the heavy atoms are shown for comparison. It can be seen how the voids (shown by the arrows) in the structure at 0 GPa are larger than those at 5.8 GPa.

Apart from the shortening of the N-H…O2 hydrogen bonds driven by the closure of the voids, there are other intermolecular interactions which exhibit significant modifications on pressure increase. At 5.8 GPa C-H…O hydrogen bond interactions are formed, which were either very weak or absent at ambient conditions. These interactions are the C2…O3 [3.04(6) Å], which connects molecules within a subunit, and C3…O2 [3.18(7) Å] and C3…O3 [3.18(6) Å], which inter-connects subunits. Weak C-H…O hydrogen bonds have C…O distances which fall within the range of 3.0-4.0 Å (Desiraju *et al.*, 1999) and, therefore, despite being initially weak, the C-H…O interactions in L-α-aspartic acid become significant at high pressure.

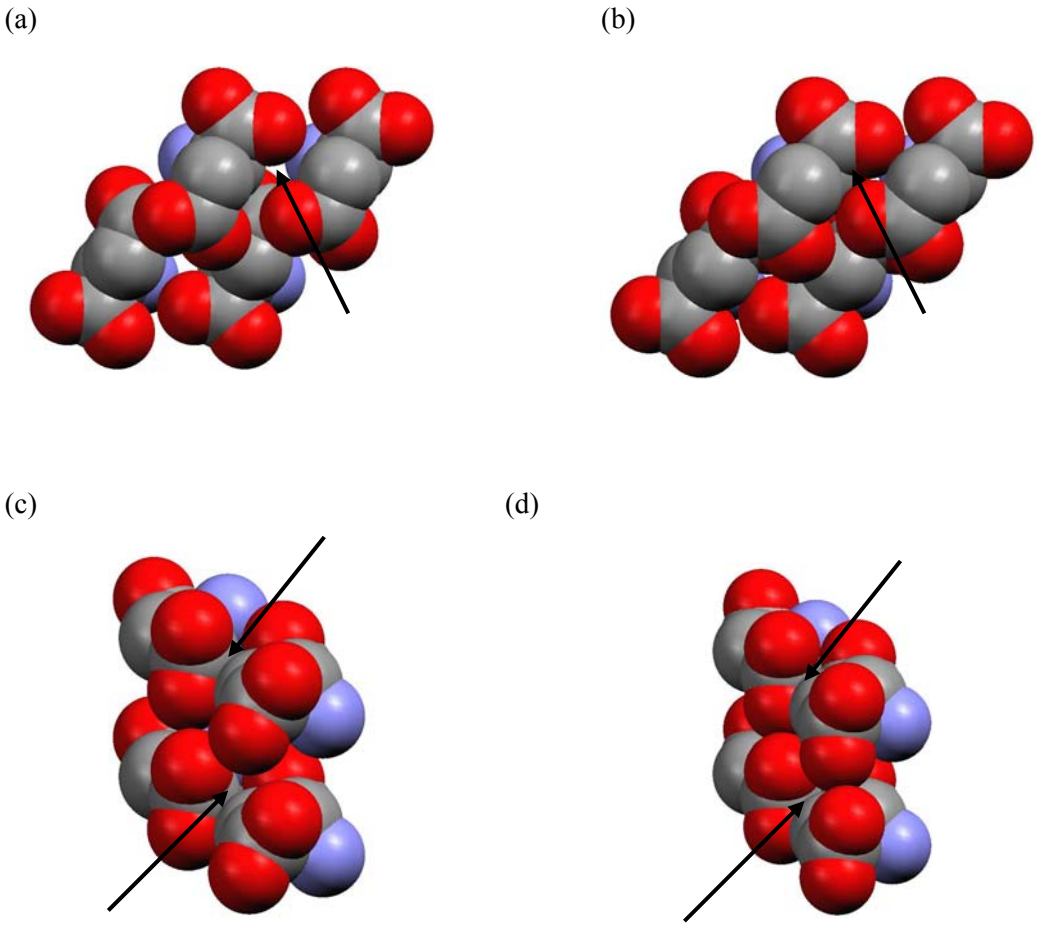


Figure 5.9: Selected space-filling representations of two R-type graph sets which occur in the crystal structure of L-α-aspartic acid at 0 and 5.8 GPa. (a, b) $R_3^4(20)$, (c, d) $R_3^3(14)$ ring motifs. Only the heavy atoms are shown for comparison. It can be seen how the voids (shown by the arrows) in the structure at 0 GPa are larger than those at 5.8 GPa.

5.2.3.4 The effect of pressure on the lattice parameters

As the unit cell volume decreases with pressure, there is an anisotropic variation of the lattice parameters as the hydrogen bonds of the underlying layered molecular structure respond to the compression (Table 5.4). From 0 to 5.8 GPa, the *a* cell dimension shortened by 0.24 Å (3.2%), the *b* lattice parameter decreased its length by around 0.90 Å (12.9%), whereas the *c* axis changed by 0.13 Å (2.5%). Accompanying these changes, the β angle increases by approximately 4°, and, overall, the cell volume decreases by 51 Å^3 (18.9%) from its original value (Figure 5.10). The b axis exhibits the greatest compression and this is principally due to the N1…O2 and N1…O1 hydrogen bonds, which lie parallel to this axis and show the most significant shortening of all the hydrogen bonds in the crystal structure.

Pressure (GPa)	***a* (Å)**	***b* (Å)**	***c* (Å)**	**β (°)**	***V* (Å^3)**
Derissen *et al.* (1968)	7.6170(10)	6.9820(10)	5.1420(10)	99.84(2)	269.44(8)
0	7.6000(17)	6.9634(17)	5.135(3)	99.76(4)	267.82(2)
1.4	7.5297(12)	6.7387(13)	5.0619(17)	101.96(3)	251.27(11)
2.9	7.462(2)	6.503(2)	5.041(3)	102.84(6)	238.50(19)
3.8	7.4214(9)	6.2863(9)	5.0249(14)	103.34(2)	228.10(8)
5.8	7.364(3)	6.059(2)	5.002(4)	103.42(6)	217.1(2)

Table 5.4: Values of the lattice parameters and volume of L-α-aspartic acid at increasing pressures.

The *a* and *c* axes exhibit very similar compressions to 5.8 GPa and they are both much less compressible than the b-axis. The relative stiffness of these axes is directly related to the N1-H…O3 and O4-H…O1 interactions, which lie parallel to the *a* axis and c-axis respectively and are the shortest hydrogen bonds in the structure.

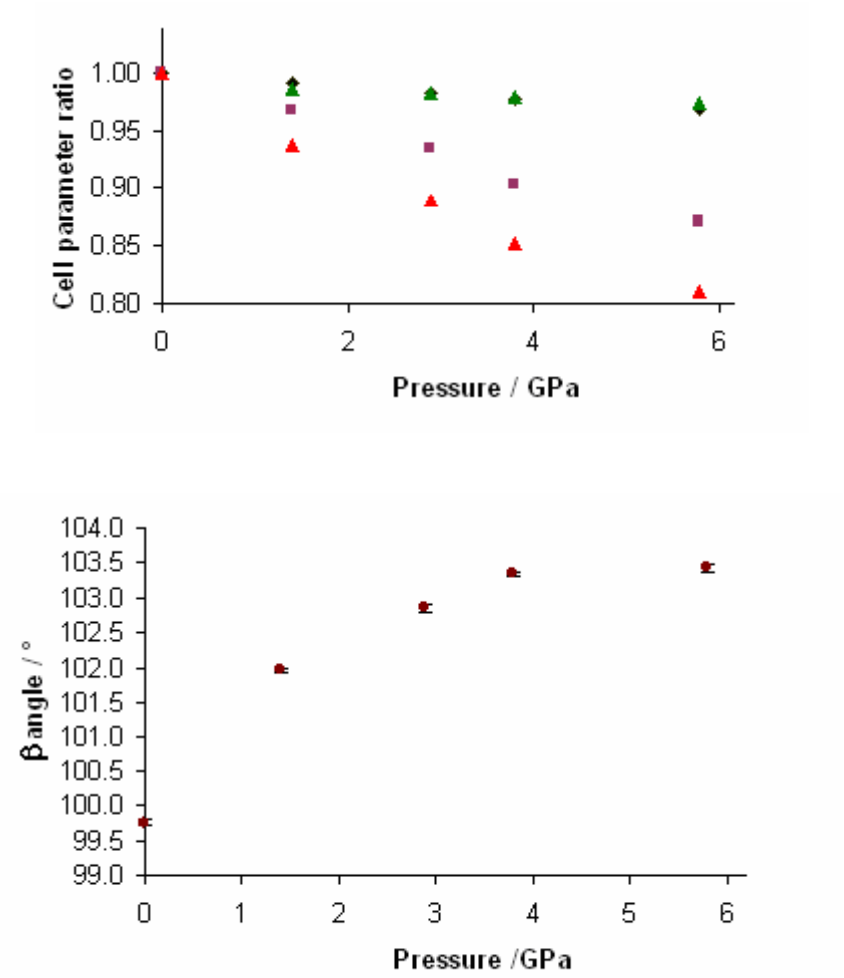


Figure 5.10: Fractional changes in the lattice parameters and cell volume of L-α-aspartic acid as a function of pressure [♦-(a/a_0)I, ■-(b/b_0), ▲-(c/c_0), ▲-(V/V_0), ✹-(β)].

5.2.4 Comparison of the ambient pressure and high pressure crystal structures of L-α-aspartic acid

5.2.4.1 Hirshfeld Surfaces

The program *Crystal Explorer* makes use of Hirshfeld surfaces to partition crystal space in molecular crystals so that the packing modes and intermolecular interactions can be examined (McKinnon *et al.*, 2004). We have used this program to

visualise the packing behaviour in the crystal structure of L-α-aspartic acid at 0 and 5.8 GPa, in order to make a more detailed comparison between them.

Hirshfeld surfaces (Hirshfeld, 1977) for the structure of L-α-aspartic acid at 0 and 5.8 GPa are shown in Figure 5.11, with the hydrogen atoms of the NH_3^+ group at the top pointing towards the viewer in each case. All the hydrogen atoms of the ammonium group actively participate in the formation of hydrogen bonding, donating the hydrogen atoms to the different oxygen atoms present in each molecule. This is shown by the orange-red region on the d_e surface (where d_e is the distance to the nearest atom centre exterior to the surface) adjacent to the oxygen and nitrogen atoms. The curvedness and shapes of the molecules at 0 and 5.8 GPa are shown in Figure 5.11. The curvedness of the molecule is a measurement of how much "shape" the surface exhibits (i.e., relatively flat areas of the surface have a low curvedness, whereas regions of the surface with rapidly varying curvature have a high curvedness). By examining the Hirshfeld surfaces at 0 and 5.8 GPa, some differences are apparent. For example, the central part of the molecule (labelled 5 in Figure 5.11), which corresponds to the NH_3^+ group, exhibits more curvedness at high pressure than at ambient pressure. This is due to the increased number of contacts with pressure due to the close proximity of the molecules. This is supported by the shape index representations of the molecules. The shape index measures the "shape" of the surface, which gives information about how molecules pack in the crystal. In these figures, contacts are represented by red areas whereas blue is for non contacts. However, the d_e representation is the type of Hirshfeld surface in which the structural changes become more apparent. In the d_e representation the external distance to the surface is measured, giving information about short contacts involving the molecules. There are important differences in the d_e surfaces for the L-α-aspartic acid molecules at 0 and 5.8 GPa. The most striking difference is the appearance of red regions (i.e. contacts) in the centre of the surface (labelled 5 in Figure 5.11) formed by the NH_3^+ group, which is due to the N1…O1 interaction, and the change in colour of yellow areas towards red regions shows that existing contacts are becoming significantly shorter (labels 1-7' in Figure 5.11) or that new contacts are forming (e.g. O4…O3, labelled 7' in Figure 5.11).

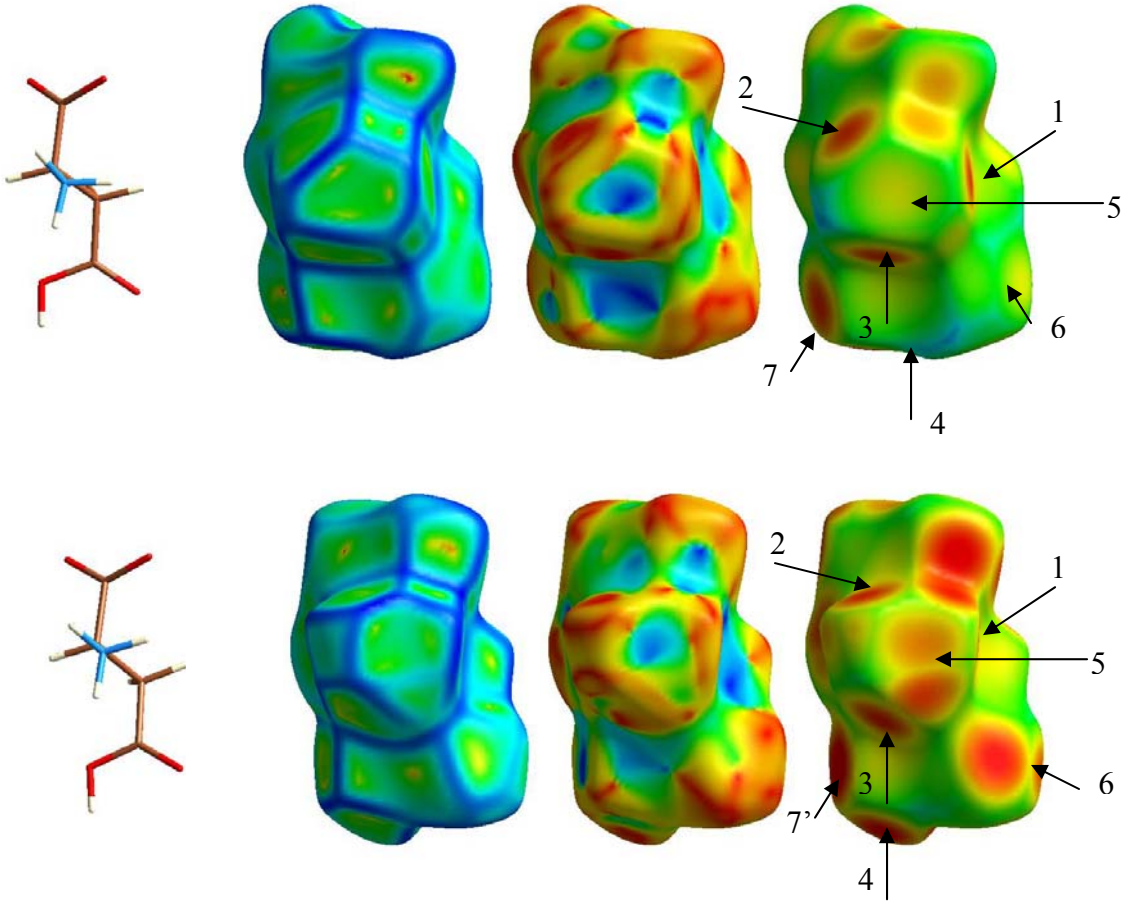


Figure 5.11: Hirshfeld surfaces for the ambient pressure (top) and high-pressure (bottom) structures of L-α-aspartic acid. Each molecule is shown with the Hirshfeld surface mapped with curvedness (left), shape index (centre) and d_e [right; for this series mapped between 1.0 (red) and 2.5 Å (blue)], where d_e is the distance to the nearest atom centre exterior to the surface. The different interactions are labelled 1-6, as shown in the text: (1) N1…O2^{i}, (2) N1…O2^{ii}, (3) N1…O3, (4) O4…O1, (5) N1…O1, (6) O3…O4, (7) O3…O1 and (7') O4…O3.

The two-dimensional fingerprint plots (plot of d_i *vs* d_e, where d_i is the distance to the nearest atom centre interior to the surface) for L-α-aspartic acid at 0 and 5.8 GPa emphasise the systematic differences between the two structures (Figure 5.12). One of the main differences is that the voids (upper region of the plots in Figure 5.12) are more compact (or less diffuse) in the high-pressure crystal structure than in the ambient pressure structure indicating that the packing is more efficient.

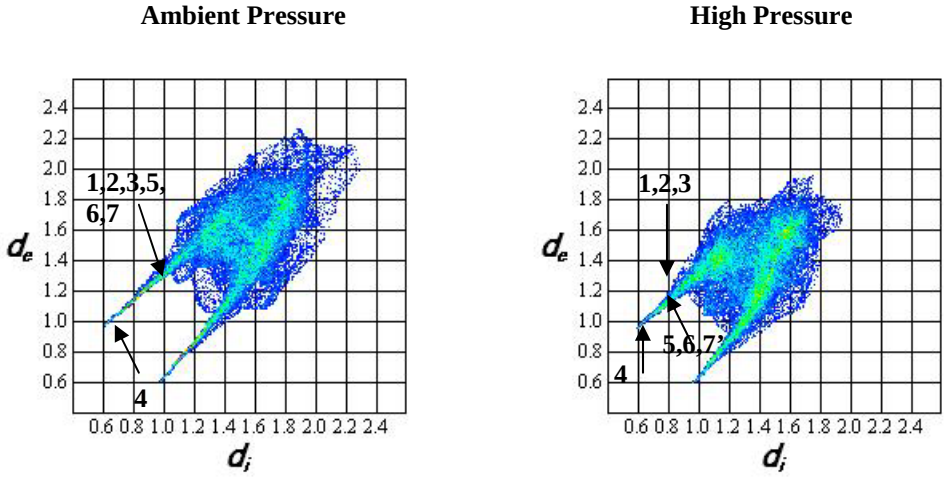


Fig. 5.12: Two-dimensional fingerprint plots for the ambient pressure (left) and high-pressure (right) structures of L-α-aspartic acid, where d_i, similarly to d_e, is the distance to the nearest atom centre interior to the surface. The different interactions are labelled 1-6, as shown in the text: (1) N1…O2^{i}, (2) N1…O2^{ii}, (3) N1…O3, (4) O4…O1, (5) N1…O1, (6) O3…O4, (7) O3…O1 and (7') O4…O3.

These fingerprint plots also give an indication of the topology of the hydrogen bonds present in the crystal structure. Every molecule of L-α-aspartic acid forms four hydrogen bonds with neighbouring molecules at ambient pressure, and only one of them involves the OH group. The three hydrogen bonds formed by the NH_3^+ group exhibit N…O distances between 2.7 and 2.9 Å, compared with the 2.5 Å length of the O4…O1 hydrogen bond formed by the OH group. Consequently, the three N…O hydrogen bonds (labelled 1, 2 and 3 on increasing N…O distance, in Figure 5.12) are found, as expected, further away from the O4…O1 hydrogen bond (labelled 4 in Figure 5.12). However, the N…O distances dramatically decrease in some hydrogen bonds when pressure is increased to 5.8 GPa, to distances of around 2.6 and 2.7 Å (labelled 1, 2 and 3 on increasing N…O distance, in Figure 5.12). The general shortening of the interatomic contacts results in the fingerprint plot moving towards the origin. The fingerprint plots also show that additional hydrogen bonds are formed at 5.8 GPa: i.e., a new N….O interaction arises between the NH_3^+ and OH groups, with a distance close to 2.9 Å (labelled 5 in Figure 5.12). These N…O interactions are

represented in different regions of the fingerprint plot, compared to that at ambient pressure, in which the N….O interactions were clustered around the same zone. Nevertheless, in both fingerprint plots the strong hydrogen bonds are represented by the presence of two spikes; one for the donor atoms and the other for the acceptor atom. It can be seen how the lengths of the spikes are shorter in the fingerprint at 5.8 GPa, as the length of the spike is directly related to the hydrogen bond distance. These spikes are representative of the O4…O1 (labelled 4 in Figure 5.12) strong interactions present at 0 and 5.8 GPa, but also of the N1…O1 strong interaction formed only at high pressure (labelled 5 in Figure 5.12). In addition to these interactions there is a O4…O3 weak interaction (labelled 6 in Figure 5.12), which strengthens with pressure (i.e., becomes shorter). An additional characteristic of the structure of L-α-aspartic acid at 0 GPa is the presence of C….O interactions, which increase their number on pressure increase. This can be seen in the area between the spikes, which is more compact in the fingerprint at high pressure than that at ambient pressure. Finally, the fingerprint plots become more symmetrical with pressure, due to the closing up of the voids.

5.3 Computational study of L-α-aspartic acid

5.3.1 Introduction

The computational method that was used to investigate the crystal structure of α-glycine will be used to study the crystal structure of L-α-aspartic acid in the present chapter. The purpose of this investigation is to extend our study to include a different type of amino acid. L-α-aspartic acid ($C_4H_7NO_4$) [see Figure 5.13(b)] is an acidic amino acid, that is it has two carboxylic acid groups (only one of the groups is charged) instead of a charged carboxylic acid group and a neutral amide group (L-α-glutamine) or a hydrogen atom (α-glycine). These carboxylic acid groups give rise to N^+-H…$^-$O and O-H…$^-$O type intermolecular interactions, which form the overall hydrogen bond network. In the previous chapter, the compressibility studie of α-glycine was discussed in detail, exploring the effect of pressure on the most relevant structural features of its crystal structures, and relating the experimental results to

those calculated using quantum mechanical methods (i.e., PW-DFT). A similar discussion will be carried out for the crystal structure of L-α-aspartic acid to continue exploring the behaviour of amino acid structure under conditions of pressure.

One of the most important features of the crystal structure of L-α-aspartic acid is the existence of an O-H…$^{-}$O hydrogen bond. This type of hydrogen bond has been classified in the literature as a strong interaction owing to its quasi-covalent nature (Steiner, 2002). The O-H…$^{-}$O interaction present in the structure of L-α-aspartic acid is formed as a combination of the acid ($RCOO^{-}$) with its conjugate base ($RCOO^{-}$). This type of combination has been reported to give rise to a strong interaction (Hibbert *et al.*, 1990), which Gilli et al. have called "negative charge assisted hydrogen bonds" (Gilli *et al.*, 1994, 2000). Steiner reported a mean O…$^{-}$O distance of 2.544(3) Å, based on a sample size of 421 compounds (Steiner, 2002). In the first part of the chapter, the experimental O…$^{-}$O distance was found to be of 2.55(2) Å, which is clearly consistent with Steiner's sample group. This interaction is around 0.2 Å shorter than the N-H^{+}…$^{-}$O interactions present in the crystal structure of L-α-aspartic acid and, as the interaction energy is correlated to its geometry, the calculated energy of the O-H…$^{-}$O hydrogen bond is expected to be higher than those values obtained for N-H^{+}…$^{-}$O interactions. During the compressibility study of the L-α-aspartic acid crystal structure, the O-H…$^{-}$O interaction exhibited the smallest compression among all the interactions that form the structure, which would appear to support our expectations. In this chapter the energies of both types of interactions, N-H^{+}…$^{-}$O and O-H…$^{-}$O bonds, together with the sublimation, lattice and proton transfer energies will be calculated and studied for comparison with previously studied amino acids.

5.3.2 Results and Discussion

The principal aims of this chapter are to not only demonstrate that PW-DFT calculations can achieve reliable estimates for sublimation, lattice and proton transfer enthalpies as well as energies of individual N-H…O hydrogen bonds present in the crystal structures of amino acids, but also to study a second type of strong interactions, which are O-H…O hydrogen bonds.

5.3.2.1 Crystal structure of L-α-aspartic acid at ambient pressure

The starting model to carry out a full quantum mechanical geometry optimisation for L-α-aspartic acid was taken from the ambient pressure X-ray diffraction study described in the previous part of this chapter, which is consistent with the X-ray study performed by Derissen *et al.* (1968) [Figure 5.13(a)]. The crystal structure is characterised by the formation of subunits of L-α-aspartic acid molecules in a head-to-head arrangement, whereas subunits are inter-connected by tail-to-tail interactions (see Figure 5.1). These subunits give rise to layers of molecules arranged parallel to the (1 0 1) Miller planes and stacked along the c axis. The structure is stabilised by three N^{+}-H…$^{-}$O=C hydrogen bonds, formed by the NH_3^{+} group, and a COO-H…$^{-}$O=C strong interaction, formed by the OH group present in each L-α-aspartic acid molecule [see Figure 5.3(a-d)]. Note that as the experimental study was part of a high-pressure crystallographic investigation, data were collected in the presence of a diamond anvil pressure cell, which consequently restricted the quantity and quality of diffraction data collected. This had repercussions on the resulting structure refinement, with the hydrogen atoms requiring to be placed geometrically, that is their locations were undetermined by the experimental data.

(a) (b)

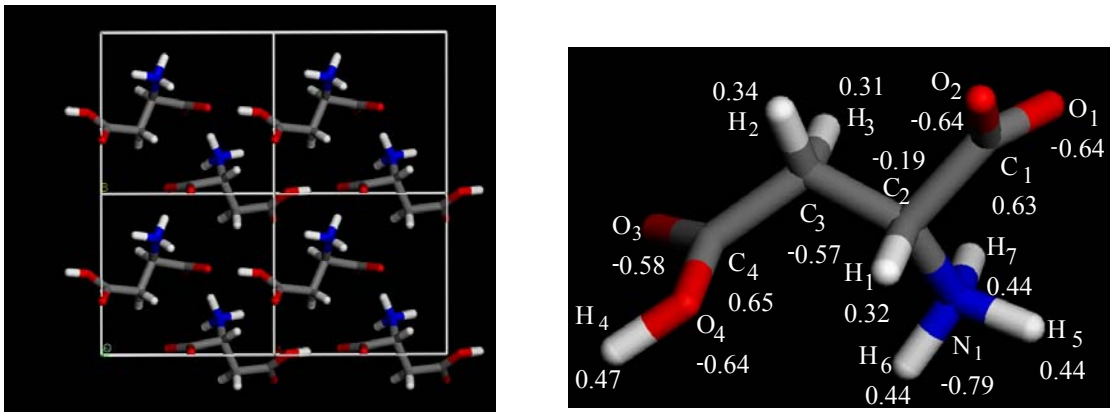


Figure 5.13: (a) The crystal structure of L-α-aspartic acid. (b) The calculated Mulliken charges for the atoms forming the L-α-aspartic acid molecule in the solid state structure.

5.3.2.2 Literature energies

All the data available from the literature relate to the sublimation, lattice and proton transfer energies of the L-α-aspartic acid crystal structure. Gaffney *et al.* (1977) reported an experimental value of 23(1) kcal mol^{-1} (~96 kJ mol^{-1}) for the sublimation energy of the structure, which was measured using mass spectrometry techniques. In terms of lattice energy, Momany *et al.* (1974) reported a value of 35.1 kcal mol^{-1} (146.7 kJ mol^{-1}), which was calculated using semi-empirical CNDO/2(ON) calculations (where ON means that the charges were established on the O and N-atoms). Nevertheless, the value for the lattice energy seems too low and therefore unlikely, based on the assumption that the sublimation energy added to the proton transfer energy should be equal to the lattice energy. Finally, Gaffney *et al.* (1977) published a calculated value of 34.3 kcal mol^{-1} (143.4 kJ mol^{-1}) for the proton transfer energy of the L-α-aspartic acid molecule.

5.3.2.3 Geometry optimisation

The results obtained by PW-DFT calculations for the geometry optimisation of the L-α-aspartic acid structure can be found in Table 5.5, together with the experimental data for comparison. From this it can be seen that in general the computational results are in good agreement with the experimental structure. The differences in cell parameters are just 0.7, 4.1, 1.7, and 3.6% for the 'a', 'b' and 'c' directions and the 'β' angle, respectively. The cell volume increased by 7.6% with respect to the experimental value. The simulation reported heavy atom positions largely consistent with the input experimental ones, with differences in the bonding parameters of the order of picometers [e.g. r(C-N)exp = 1.489(9) Å, r(C-N)calc = 1.489 Å]. The locations of the hydrogen atoms, placed geometrically in the experiment, were up-dated quantum mechanically and were found to remain in essentially equivalent positions, with any differences noted also of the order of picometers.

Parameters	**Experimental**	**Calculated**		
	Solid	Solid	Supercell (not relaxed)	Supercell (relaxed)
Lattice [Å, °]				
a	7.6000(17)	7.6538	16	16
b	6.9634(17)	7.2486	16	16
c	5.135(3)	5.2236	18	18
α=γ, β	90, 99.76(4)	90, 96.18	90, 96.18	90, 96.18
Z	2	2	1	1
Volume [Å^3]	267.82(16)	288.12		
Space/point group	*P12$_1$1*	*P12$_1$1*	*P1*	*P1*
Geometry [Å, °]				
Dimer 1				
rN1…O3	2.78(2)	2.813	—	—
rN1-H12	1.003	1.054	1.054	—
rO3…H12	1.799	1.766	—	—
∠H12-N1-H11	109.6	107.0	107.0	—
∠H12-N1…O3	167.8	171.7	—	—
Dimer 2				
rN1…O2	2.82(3)	2.754	—	—
rN1-H11	0.995	1.056	1.056	1.021
rO2…H11	1.873	1.699	—	—
∠H11-N1-H13	109.5	107.2	107.2	—
∠H11-N1…O2	158.5	176.3	—	—
Dimer 3				
rN1…O2	2.90(4)	2.892	—	—
rN1-H13	1.004	1.046	1.046	1.025
rO2…H13	1.933	1.860	—	—
∠H13-N1-H12	108.9	109.6	109.6	
∠H13-N1…O2	160.8	168.5	—	—
Dimer 4				
rO4…O1	2.55(2)	2.534	—	—
rO4-H1	0.900	1.052	1.052	0.981
rO1-H12	—	—	—	1.011
rO1…H1	1.647	1.483	—	—
∠H1-O4…O1	179.6	176.1	—	—
Energy [kcal mol^{-1}]	—			
Lattice	35.1*	—	69.4	—
Sublimation	23(1)**	—	—	35.1
Proton Transfer	34.3***	34.3	—	—
Total energy [per molecule, eV]	—	-2762.04818	-2759.041984	-2760.52573

Table 5.5: Comparison between the experimental and calculated (PW-DFT) structures for the crystal structure of L-α-aspartic acid (* from Momany *et al.*, 1974; **, *** from Gaffney *et al.*, 1977).

5.3.2.4 Calculated energies

The calculated lattice energy found in the literature (Momany *et al.*, 1974), of 35.1 kcal mol^{-1} (146.7 kJ mol^{-1}), was obtained through semi-empirical CNDO calculations. However, on the basis of experimentally reported sublimation energy and proton transfer energies, this lattice energy value seems rather low. Our results showed that the calculated lattice energy was found to be 69.4 kcal mol^{-1} (290.1 kJ mol^{-1}), which is much closer to the expected value for this property. Additionally, the value calculated for the lattice energy of glycine (57.1 kcal mol^{-1} or 238.7 kJ mol^{-1}) is very close to the result obtained for aspartic acid, differing significantly from the calculated lattice energy reported by Momany *et al.* (1974). Although no direct comparison of these two values can be done due to the differences in the approach to perform the calculations, it is known that semi-empirical methods are not as accurate as first-principles methods. The magnitudes of sublimation energies for amino acids depend greatly on the method employed. The values obtained from mass spectra are much smaller than those obtained from Knudsen effusion experiments (No *et al.*, 1994). No *et al.* explained how this is merely due to the fact that mass spectra data do not represent the following trend: most of the sublimation energy arises from the dispersion energy and is linearly dependent on the molecular polarizability, which is proportional to molecular size. The calculated value for the sublimation energy of the L-α-aspartic acid crystal structure falls within the range of experimental energy values found for other amino acids, such as α-glycine [32.6(1) kcal mol^{-1}], alanine [33.0(2) kcal mol^{-1}] and valine [38.9(2) kcal mol^{-1}] (Svec *et al.*, 1965). Finally, the proton transfer energy calculated from the difference between the sublimation and lattice energies was found to be 34.3 kcal mol^{-1}, which is in close agreement with the 39.7 kcal mol^{-1} (166.0 kJ mol^{-1}) obtained by Gaffney *et al.* (1977) [26.1 kcal mol^{-1} or 109.1 kJ mol^{-1}] during this study, and with other results for amino acids reported in the literature, such as cysteine [26.9 kcal mol^{-1} or 112.4 kJ mol^{-1}] and serine [24.3 kcal mol^{-1} or 101.6 kJ mol^{-1}] (No *et al.*, 1994). Now that the values for the calculated and experimental sublimation and proton transfer energies are known, the calculated and experimental values for the lattice energy can be discussed. The lattice energy of a system is equal to the sum of the sublimation and proton transfer energies of that system. Therefore, by adding up the experimental

values for both the sublimation (23(1) kcal mol^{-1}; Gaffney *et al.*, 1977) and proton transfer (39.7 kcal mol^{-1}; Gaffney *et al.*, 1977) energies, we would expect the lattice energy to be of around 63 kcal mol-1 (263 kJ mol-1), which is in good agreement with the PW-DFT calculated lattice energy of 69.4 kcal mol^{-1} (290.1 kJ mol^{-1}).

The lattice energy value obtained by PW-DFT can be further tested by performing an Ewald summation, taking the optimised structure and calculated Mullilken charges obtained by the quantum mechanical simulations [Figure 5.13(b)] as input for the program GULP (Gale *et al.*, 2003), as done for α-glycine in the previous chapter. The value thus obtained for L-α-aspartic acid was 64.3 kcal mol^{-1} (268.8 kJ mol^{-1}), which is only 5.1 kcal mol^{-1} (22.6 kJ mol^{-1}) lower than the value reported above by PW-DFT calculations. Thus, results show a close match between the values calculated using the Ewald sum and those obtained through DFT-PW calculations. This close agreement is prove that the crystal structure packing interactions can be described as largely electrostatic, becoming a very useful tool in the study of the solid state energetics of charged systems, such as amino acids.

In order to assign energy values to the individual hydrogen bonds present in the crystal structure of L-α-aspartic acid, four dimer models were constructed, labelled dimer models 1-4 [Figure 5.14(a-d)], to obtain each hydrogen bond in isolation. Thus, dimer 1 concerns N_1-H_5…O_2 (1.766 Å, 171.7°), which connects molecules of L-α-aspartic acid to form chains of molecules parallel to the c axis [Figure 5.14(a)], described as *C(5)* in graph set notation. Dimer 2 is formed by two molecules of L-α-aspartic acid linked by the N_1-H_6…O_3 interaction (1.699 Å, 176.3°), which links molecules to form *C(6)* chains, which are generated by the 2_1 screw axis parallel to the b axis [Figure 5.14(b)]. Dimer 3 describes the hydrogen bond N_1-H_7…O_2 (1.860 Å, 168.5°), which holds molecules together to form subunits [see Figure 5.14(c)] and forms *C(5)* chains parallel to the b axis. Dimer 4 relates to the O_4-H_4…O_1 interaction (1.483 Å, 176.1°), which forms *C(7)* chains that run parallel to the a axis and is created by the OH group of the L-α-aspartic acid molecule [Figure 5.14(d)]. Although all the interactions exhibit periodicity in the crystal structure, owing to difficulties encountered when constructing the models for the calculations (other interactions were always present together with the targeted

hydrogen bond in the selected model), one of the dimer models containing the linkage N_1-H_6…O_3 could not be separated from other dimers, and therefore this model could not be used to obtain the energy of this interaction in isolation. Rather, the value for the N_1-H_6…O_3 is inferred by subtracting the energies of all the other H-bond interactions from the lattice (i.e. total) energy. In addition to this, the construction of an aperiodic model for dimer 4, O_4-H_4…O_1, was not possible due to its alignment with respect to the axes.

We can relate the geometry of the four different hydrogen bonds present in the L-α-aspartic acid structure with their expected individual energies, as was done for α-glycine. On the basis of geometry alone, we would expect the energy of the hydrogen bond interaction to decrease as follows: dimer 4 > dimer 2 > dimer 1 > dimer 3 (Jeffrey, 1997). The same *ab initio* and classical methods as used to calculate the strengths of the individual hydrogen bonds for the α-glycine system were applied to L-α-aspartic acid. The results for these calculations can be found in Table 5.6. Thus, the PW-DFT values, at 28.6, 6.5, 13.0 and 21.2 kcal mol^{-1} (119.5, 27.2, 54.3, and 88.6 kJ mol^{-1}) for models 1-4, respectively. These results fall in the same range of energy as that found for the hydrogen bond energies in α-glycine, of energies between 1 and 30 kcal mol^{-1} (4 and 125 kJ mol^{-1}). On the basis of geometry alone, however, the calculated energies obtained for the hydrogen bonded dimers did not follow the predicted trend, instead: dimer 1 (28.6 kcal mol^{-1} or 119.5 kJ mol^{-1}) > dimer 4 (21.2 kcal mol^{-1} or 88.6 kJ mol^{-1}) > dimer 3 (13.0 kcal mol^{-1} or 54.3 kJ mol^{-1}) > dimer 2 (6.5 kcal mol^{-1} or 27.2 kJ mol^{-1}). The discrepancy between the calculated results and those to be expected is mainly due to secondary interactions arising from the interaction between atoms of different molecules. These secondary interactions have been shown to be of great significance in the hydrogen bonding present in the α-glycine crystal structure (see Chapter 4). We will go back to this point later, when discussing the electrostatic contribution to the hydrogen bonding.

In order to study the σ-bond cooperativity effect in the crystal structure of L-α-aspartic acid, four aperiodic hydrogen bonds were constructed. Results show that the intermolecular interactions were strengthened by very different degrees (9-25%), and the range is in good agreement with what was found for α-glycine (see Table 5.6).

As the results stated above show, in the case of L-α-aspartic acid there is not a clear correlation between the individual hydrogen bond energies and the geometries of the four interactions. However, as it was shown for α-glycine, this correlation might only become apparent upon close examination of the individual atom-atom electrostatic pair contributions. Thus, we have calculated the total electrostatic contribution to the hydrogen bond energies for dimers 1-4 in the L-α-aspartic acid crystal structure, using V_Coulombic, which resulted in very similar values as obtained by Ewald summation (GULP) (see Table 5.6). When comparing these two sets of energy values with the results obtained for the individual interactions using PW-DFT, all the calculated energy values for dimers 1-4 were comparable.

The different energies calculated for the four individual hydrogen bonds present in the crystal structure of L-α-aspartic acid can be explained in terms of the electrostatic nature of these four interactions. When comparing the overall repulsive and attractive energies arising from the intermolecular atom-atom interactions between molecules in dimer models (Table 5.7), it was found that the four intermolecular interactions present repulsive energies of 1787.0, 1765.7, 1935.0 and 1909.2 kcal mol^{-1} (7469.7, 7380.6, 8088.3, and 7980.5 kJ mol^{-1}) for dimers 1 to 4 respectively, and attractive energies of -1795.3, -1791.4, -1940.0 and -1545.2 kcal mol^{-1} (-7504.3, -7488.0, -8109.2 and -6458.9 kJ mol^{-1}). In addition to this, the energy of the H…O interaction was found to be 49.9, 52.1, 49.5 and 64.9 kcal mol^{-1} (208.6, 217.8, 206.9 and 271.3 kJ mol^{-1}) for dimers 1-4, respectively. The calculated electrostatic energy values obtained for the different H…O interactions show how the H…$^{-}$O interactions of the N^{+}-H…$^{-}$O hydrogen bonds are very much alike in strength (~ 50 kcal mol^{-1}), whereas the one in O-H…$^{-}$O is almost 15 kcal mol^{-1} stronger. This was to be expected, since the electrostatic energy only depends on the charges of the atoms involved in the interaction and the interatomic distance between them. From these result, it can be seen how the strength of the O-H…$^{-}$O interaction is stronger

than the N^{+}-H…$^{-}$O hydrogen bonds, owing its strength to its quasi-covalent character. The graphic representation of the atom-atom electrostatic interactions which arise from the dimer interactions are shown in Figure 5.15.

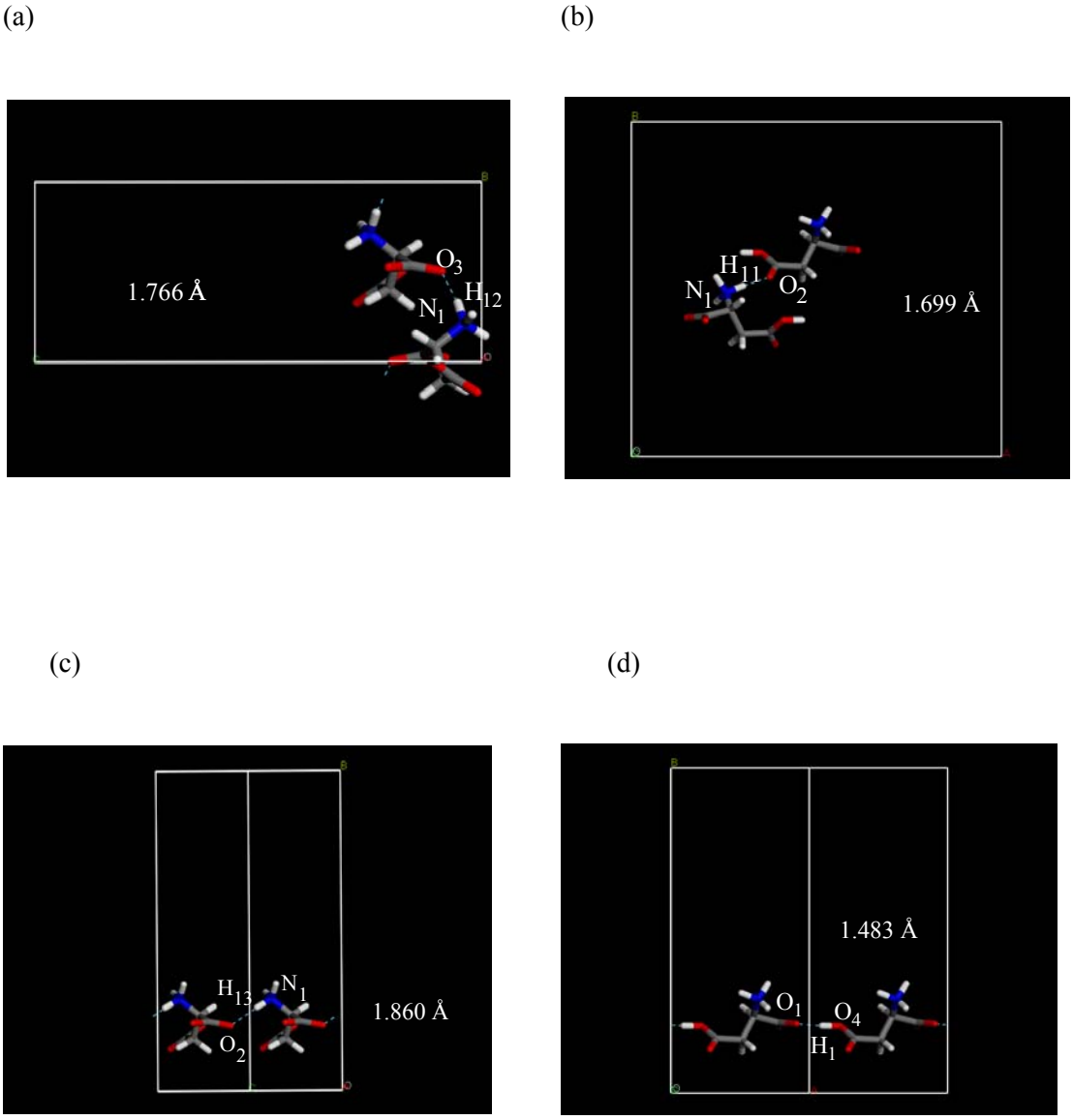


Figure 5.14: The four dimer models used in the calculation of the individual hydrogen bond energies (dimers 1-4, respectively). Note the bond order in models 1, 3-4 is two, whereas for model 2 it is one.

Finally, no simulations were run using the GAUSSIAN98 package, since it was already shown in previous chapters how the energies calculated using GAUSSIAN98 and PW-DFT are in good agreement.

Model			V_Coulombic	CASTEP		GULP	
	r(D-H…A) *r* (D…A)	∠(D-H…A) ∠(C-D…A)	Energy/ H- bond	Energy/ eV	Energy/ H- bond	Energy/ eV	Energy/ H- bond
Dimer 1	1.766	171.7					
Periodic	2.813	104.7		-2760.28357	28.6	—	
Aperiodic			25.4	-2759.61679	26.5	-1.13628	26.2
Dimer 2	1.699	176.3					
Periodic	2.754	112.9	—	—	6.5[b]	—	
Aperiodic			8.3	-2759.14907	4.9	-0.40960	9.4
Dimer 3	1.860	168.5					
Periodic	2.892	108.3	—	-2759.60552	13.0	—	
Aperiodic			4.9	-2759.25573	9.9	-0.29488	6.8
Dimer 4	1.483	176.1					
Periodic	2.534	115.1	—	-2759.96136	21.2	—	
Aperiodic			16.9	-2759.45277	19.0	-0.79221	18.3
Monomer	—	—	—	-2759.04198	—	—	—

[a] (Volkov *et al.*, 2003)
[b] Result not calculated directly, but rather inferred from the summation of other results, with respect to the calculated lattice energy.

Table 5.6: Energies obtained per molecule for the L-α-aspartic acid dimer and monomer models, and the resulting energies per hydrogen bond (kcal mol^{-1}) (D – H … A, C-D…A; *r*/Å, ∠/°, D = donor, A = acceptor, C = Carbon atom).

V_Coulombic	**Energy/ kcal mol^{-1}**			
Energy Type	*Dimer 1*	*Dimer 2*	*Dimer 3*	*Dimer 4*
Attractive	-1791.4	-1795.3	-1939.9	-1545.2
Repulsive	1765.7	1787.0	1935.0	1528.3
Total (hydrogen bond)	-25.7	-8.3	-4.9	-16.9
H…O interaction	-52.1	-49.9	-49.5	-64.9

Table 5.7: Calculated total, attractive and repulsive electrostatic energies (kcal mol^{-1}) for the interactions present in dimers 1-4 in the crystal structure of L-α-aspartic acid. Additionally, the energy of the H…O interaction (kcal mol^{-1}) can also be found.

(a) (b) (c) (d)

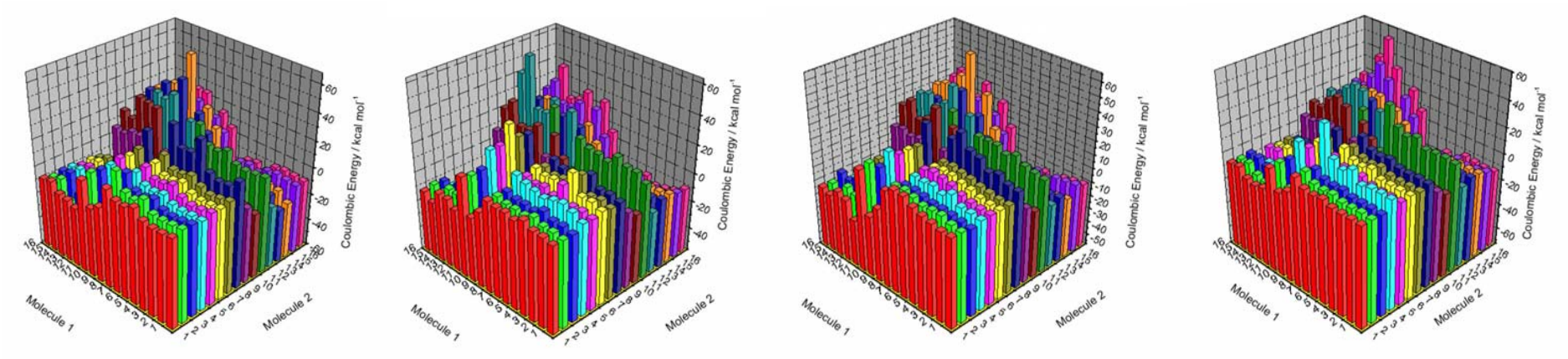


Figure 5.15: Histogram of the atom-atom contributions to the total electrostatic interaction energy (kcal mol^{-1}) marked according to atom type for (a) model 1, (b) model 2, (c) model 3 and (d) model 4 in L-α-aspartic acid.

5.3.3 Compressibility of the L-α-aspartic acid crystal structure according to computational results

The experimental compressibility study of the L-α-aspartic acid crystal structure was described in detail in the first part of this chapter. When going from 0 to 5.8 GPa, the unit cell parameters decreased their length from 7.6000(17), 6.9634(17) and 5.135(3) Å for the 'a', 'b' and 'c' cell dimensions respectively, to 7.364(3), 6.059(2) and 5.002(4) Å, with changes of 3% along the 'a' and 'c' directions, and 13% along the 'b' direction. The overall change in the unit cell volume was of 19%, decreasing from 267.82(16) to 217.1(2) $Å^3$. The β angle increased its value from 99.76(4) to 103.42(6)°. The changes in the cell parameters are a direct consequence of the shortening of the intermolecular interactions present in the crystal structure (i.e. dimers 1-4) upon pressure. Experimental results showed that the intermolecular interactions decreased their length by 3.6, 4.9, 6.3, and 3.1% for dimers 1-4, respectively. This means that the strength of the hydrogen bonds according to the experimental compressibility results would follow the scheme: dimer 4 > dimer 1 > dimer 2 > dimer 3. Although there is a clear correlation between the geometry of the interactions and the amount they compress, as it was shown in chapter 4 for α-glycine, when interactions have similar geometry (e.g. dimers 1 and 2) the correlation weakens. In the optimised crystal structure the compressibility of the interactions in terms of their optimised geometry would follow the following scheme: dimer 4 > dimer 2 > dimer 1 > dimer 3, which disagrees with what was found during the experimental study due to the geometry similarities between dimers 1 and 2. However, the importance of secondary interactions was also discussed in previous chapters, demonstrating that the bigger the system (i.e., more molecules in the structure), the more important these atom-atom interactions become due to the increase in the number of contacts between atoms, especially when the participating atoms are charged. In the case of L-α-aspartic acid the correlation stated above does not seem to hold and instead, changes of the intermolecular distances follow the following scheme: dimer 3 > dimer 2> dimer 1> dimer 4, where dimer 3 is the most compressible interaction and dimer 4 is the least compressible hydrogen bond. Thus, dimers 3 and 4 fulfill the expected behaviour, exhibiting the largest and lowest change

in the series, respectively. Nevertheless, in the previous section it was shown how the secondary effects due to interatomic contacts have a profound weight on the overall energy of the hydrogen bond, which was also seen in previous chapters, especially for α-glycine Thus, on the basis of the geometry of the H…O interactions, the expected changes would follow the scheme dimer 3 > dimer 2 > dimer 1 > dimer 4, which do agree with the results obtained from the experimental compressibility study. In summary, the compressibility of the intermolecular hydrogen bonds present in a crystal structure depends mainly on their geometry (i.e., distances and angles), but also on atom-atom secondary interactions which might give rise to repulsive or attractive energy, which can cancel out the energy of the H…O interactions. In addition to this, pressure has an important effect on the L-α-aspartic acid crystal structure, with the consequent shortening of the hydrogen bonds and the unit cell dimensions, but also, and as a direct consequence of those changes, with the 'close up' of the voids present in the packing. In general, these voids are due to the formation of rings by four strong hydrogen bonds (dimers 1-4) together with two weaker interactions (N1…O1 and O4…O3), which coexist to form the crystal structure of L-α-aspartic acid. The weight of the interactions is not equally divided among the six interactions and therefore, when pressure is applied to them they must compete for space and the geometry of the interactions might not be the most relevant factor to explain the structural changes taking place in the molecular packing. Although all the efforts were made to fully understand the behaviour of the hydrogen bonds under pressure, trying to relate their energies to their compressibility, there was not a definite conclusion on why or how these changes take place and further investigations must be carried out in the area, such as the study of the effect of other interactions on a hydrogen bond when they are in close proximity.

5.3.4 Computational

5.3.4.1 Quantum mechanical simulations

5.3.4.1.1 Crystal structure calculations

Calculations on the crystal structure of L-α-aspartic acid were performed using the CASTEP 2.2 package available through Materials Studio suite of software (Accelerys Inc., 2002). The valence electrons were modelled using a plane-wave basis set expressed at an energy cut-off of 450 eV, which was found to converge the total energies to better than 2.0 meV per unit cell in both cases. The electronic core wave function was described using the standard ultrasoft pseudopotentials available with the software package. The symmetry-reduced k-point sets used to sample reciprocal space were generated using Monkhorst-Pack grids (Monkhorst *et al.*, 1977) (dimensions 3 x 3 x 4), giving 12 k-points in the symmetry-reduced first Brillouin zone for L-α-aspartic acid. The generalised gradient functional PBE was used to model the electronic exchange and correlation.

The initial structures used to perform the geometry optimisations were taken from the ambient pressure X-ray diffraction study of L-α-aspartic acid, which was presented in the first part of the present chapter. The optimisation of the atomic positions and unit cell parameters were performed on alternate cycles using the BFGS method until the convergence criteria were met (maximum energy change per atom = 1 x 10^{-5} eV, maximum root-mean-square force = 0.03 eV $Å^{-1}$, maximum RMS stress = 0.05 GPa and maximum RMS displacement = 0.001 Å;).

A population analysis calculation was performed on the optimised structure to obtain the Mulliken charges for each atom present in the crystal lattice. Since the Mulliken population analysis is formulated in terms of atomic orbitals (Mulliken, 1955), the electronic wave function from a wave is first projected to an atomic orbitals basis set using the scheme of Sanchez-Portal (Sanchez-Portal et al., 1995).

The charge spilling parameter for the spin component was found to be 1.17% indicating a successful charge partitioning was obtained.

5.3.4.1.2 Supercell calculations

These were performed on single, zwitterionic and effectively isolated molecules of L-α-aspartic acid, for comparison with the energy per molecule in the solid state, and thus allowing the deduction of the total intermolecular interaction energies in the two crystal structures. Zero interaction between the nearest neighbouring cells was obtained by increasing the cell size and observing the change in the total energy. A cell size of 16 x 16 x 18 Å was found to break all intermolecular interactions, without giving rise to overly long computational times. Apart from a reduction in the number of k-points (to one, the gamma point), all other specifications for these calculations were the same as those applied for the corresponding crystal structure optimisation. The energy of the supercell (no atom relaxation) was compared to the solid state energy of the optimised structure to obtain the lattice energy. In addition to this, an atom-only optimisation calculation was performed on a second supercell to obtain the neutral pseudo-isolated molecule; this energy was then compared to the energy of the optimised crystal structure (per molecule) in order to obtain an estimate of the sublimation energy.

5.3.4.1.3 Dimer models

These models were constructed to estimate the individual contributions of each of the intermolecular interactions towards the lattice energy. All calculations performed were single-point energies (to avoid conversion to the more stable neutral gas-phase conformation) using the same basis set cut-off and convergence criteria as reported for the full crystal structure optimisations; the k-point sampling grid was reduced to the gamma point.

All the calculations carried out during this computational study were rigorously tested to ensure that the basis sets had reached an acceptable level of convergence.

Consequently, the comparison of energies between the supercell and the rest of the calculations (*i.e.*, crystal structure and dimer calculations) is legitimate: we estimate any error incurred due to inconsistencies in basis sets to be within 0.01 – 0.003 eV (i.e., 0.2 – 0.07 kcal mol^{-1}) for the sublimation, lattice, proton transfer and hydrogen bond energies.

5.3.4.2 Classical mechanical calculations

5.3.4.2.1 Ewald summations

Finally, to obtain the electrostatic contributions to the lattice energy (Ewald sum) the simulation package, GULP (Gale *et al.*, 2003), was used, taking the PW-DFT optimised structure and Mulliken charges as input.

5.3.4.2.2 Coulombic interaction energy calculations

This was obtained *via* an in-house program V_Coulombic, written to calculate the atom-atom pair electrostatic energy (Coulombic potential) contributions from a given set of coordinates and point charges. Input files for this program took the optimised PW-DFT crystal structure geometries and Mulliken charges.

5.4 Conclusions

We have investigated the compression of the L-α-aspartic acid at pressures between 0 and 5.8 GPa and, although no indication of a phase transition was observed, the N-H…O and O-H…O interactions exhibited considerable shortening with the N…O distance approaching the minimum distance observed for this type of interaction. Accompanying these hydrogen bond changes, the conformation of the L-α-aspartic acid molecule also became more planar. The structure was described in detail and features were compared to other amino acid structures, in order to understand the changes taking place on pressure increase on these systems. As a

consequence of the increase of pressure, new N…O and C…O interactions were formed.

Following the experimental study of the compressibility of the L-α-aspartic acid crystal structure, a computational study was carried out to complete the experimental model (i.e., locate hydrogen atom positions) and give a full picture of the crystal structure, from the geometry of the hydrogen bonding to the energetics of the systems (i.e., sublimation, lattice and proton transfer energies), as well as the energies of the individual hydrogen bond interactions. Results showed that the sublimation, lattice and proton transfer energies are in agreement with the reference values and are of the same order of magnitude than those calculated for α-glycine in Chapter 4. The calculated energies for the four hydrogen bonds present in the crystal structure of L-α-aspartic acid fell in the same range of energies (1-30 kcal mol^{-1} or 4-120 kJ mol^{-1}) as those found in the previous chapter, for α-glycine. Finally, the electrostatic contribution to the lattice and hydrogen bond energies was calculated using the Ewald sum and the Coulomb potential. These calculations show that the crystal structure of L-α-aspartic acid is dominated by electrostatic interactions as was expected owing to the zwitterionic state of the L-α-aspartic acid molecules. Hence, the secondary atom-atom electrostatic interactions were calculated and results were plotted in a histogram to show the differences among the various hydrogen bonds present in the crystal structure of L-α-aspartic acid.

In parallel to these results, different computational methods were used (quantum mechanics and molecular mechanics), such as DFT delocalised (plane wave) basis sets, Ewald sum and Coulombic potential, in order to benchmark the calculated results and to get a detailed basis for the energetics of the L-α-aspartic acid structure.

5.5 References

Angel, R. J. (2004). *J. Appl. Cryst.* **34**, 486-492.

Bernstein J.; Davis, R. E.; Shimoni, L.; Chang, N. L. (1995). *Angew. Chem. Int. Ed. Engl*. **34**, 1555-1573.

Betteridge, P. W.; Carruthers, J. R.; Cooper, R. I.; Prout, K.; Watkin, D. J. (2003). *J. Appl. Cryst.* **37**, 410-416.

Biswas, A. B.; Hughes, E. W.; Sharma, B. D. (1968). *Acta Cryst.* **B24**, 40-50.

Blessing, R. H. (1987). *Cryst. Rev.* **1**, 3-58.

Blessing, R. H. (1989). *J. Appl. Cryst.* **22**, 396-397.

Bruker-AXS (1999). GEMINI. version 1.01. Bruker-AXS, Madison, Wisconsin, USA.

Bruker-AXS (2003). SAINT version 7. Bruker-AXS, Madison, Wisconsin, USA.

Bruker-AXS (1997-2001). *SMART*. version 5.049-5.059. Bruker-AXS, Madison, Wisconsin, USA.

Dawson, A.; Allan, D. R.; Clark, S. J.; Parsons, S.; Ruf, M. (2004). *J. Appl. Cryst.* **37**, 410-416.

Derissen, J. L.; Endeman, H. J.; Peerdeman, A. F. (1968). *Acta Cryst.* **B24**, 1349-1354.

Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology.* Oxford University Press, New York, pp. 350-363.

Etter, M. C.; MacDonald, J. C.; Bernstein, J. (1990). *Acta Cryst.* **B46**, 256-262.

Gaffney, J. S.; Pierce, R. C.; Friedman, L. (1977). *J. Am. Chem. Soc.* **99**, 4293-4298.

Gale, J. D. & Rohl, A. L. (2003). *Molecular Simulation*, **29**, 291-341.

Jeffrey, G. A. (1997). *An introduction to hydrogen bonding*, Oxford University Press, UK.

Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. (1994). *J. Am. Chem. Soc.* **116**, 909-915.

Gilli, G. & Gilli, P. (2000). *J. Mol. Str.* **552**, 1-15.

Hibbert et al., 1990

Hirshfeld, F. L. (1977). *Theor. Chim. Acta*. **44**, 129.

Iitaka, Y. (1960). *Acta Cryst*. **13**, 35.

Llamas-Saiz, A. L.; Foces-Foces, C.; Elguero, J.; Meutermans, W. (1994). *Supramol. Chem.* **4**, 53.

Marsh, R. E. (1958). *Acta Cryst.* **11**, 654.

Materials Studio (2001-2005).Accelerys Software Inc.

McKinnon, J. J.;Spackman, M. A.; Mitchell, A. S. (2004). *Acta Cryst.* **B60**, 27.

Merrill, L. & Bassett, W. A. (1974). *Rev. Sci. Instrum.* **45**, 290-294.

Moggach, S. A.; Allan, D. R.; Lozano-Casal, P.; Parsons, S. (2005). *J. Synchrotron Rad.* **12**, 590-597.

Moggach, S. A; Allan, D. R.; Morrison, C. A.; Parsons, S. & Sawyer, L. (2005). *Acta Cryst.* **B61**, 58-68.

Momany, F. A.; Carruthers, L. M.; Scheraga, H. A. (1974). *J. Phys. Chem.* **78**, 1621-1630.

Monkhorst, H. J. & Pack, J. D. (1977). *Phys. Rev. B*, **13**, 5188-5192.

Mulliken, R. S. (1955). *J. Chem. Phys.* **23**, 1833.

No, K. T.; Cho, K. H.; Kwon, O. Y.; Jhon, M. S.; Sheraga, H. A. (1994). *J. Phys. Chem.* **98**, 10742-10749.

Sanchez-Portal, D.; Artacho, E.; Soler, J. M. (1995). *Solid State Commun.*, **95**, 685.

Steiner, T. (2002). *Angew. Chem. Int. Ed.* **41**, 48-76.

Suresh, C. G.; Ramaswamy, J.; Vijayan, M. (1986). *Acta Cryst.* **B42**, 473-478.

Svec, H. J. & Clyde, D. D. (1965). *J. Chem. Eng. Data*, **10**, 151-152.

Chapter 6

Structural and Computational Study of L-α-glutamine

6.1 Introduction

L-α-glutamine is a protein amino acid and is found in all forms of life. It is classified as a semi-essential or conditionally essential amino acid. L-α-glutamine is very versatile, participating in many reactions in the body and, for example, is important in the regulation of acid-base balance. L-α-glutamine participates in the formation of purine and pyrimidine nucleotides, amino sugars (such as glucosamine and L-glutamate) and other amino acids (for example nicotinamide, adenine, dinucleotide and glutathione). It also participates in protein synthesis, energy production and the formation of D-glucose and glycogen. Importantly, L-α-glutamine can serve as the primary respiratory substrate for the production of energy in enterocytes and lymphocytes. It is considered to be an immunonutrient, and supplemental L-α-glutamine is used in medical foods for such stress situations as trauma, cancer, infections and burns (Skubitz *et al.*, 1996; Anderson *et al.*, 1998).

The first structural studies performed on L-α-glutamine were published by Cochran et al. (1952) and Koetzel *et al.* (1973), who reported that *"the structure is stabilized by a three-dimensional network of N-H…O hydrogen bonds … There are five unique hydrogen bonds, one for each hydrogen attached to nitrogen … The five N-H…O bonds are all significantly bent and are distributed over five neighbouring molecules, resulting in a complicated hydrogen bonding pattern"* (Koetzel *et al.*, 1973). Due to the limited description of the crystal structure of L-α-glutamine available in the literature, a more exhaustive study of the molecular packing was performed during the production of the work presented in this thesis, so a better description of the L-α-glutamine structure can be given in this chapter.

The work presented in this thesis is one of a number of investigations we are conducting on the effect of pressure on the crystal structures of different organic and biological compounds such as acetone, cyclopropylamine, ethanol, methanol, L-serine and L-cysteine, among others (Lozano-Casal *et al.*, 2005; Allan *et al.*, 1998, 1999, 2001; Moggach *et al.*, 2005*a*, 2005*b*). The aim of these studies is to understand how the various intermolecular and intra-molecular interactions respond to pressure, which

can often result in the formation of new high-pressure polymorphs. Indeed, the degree to which bond compressibility can be explained and the extent to which a structure can be compressed before a phase transition takes place, are profoundly important questions in the field of high-pressure structural chemistry.

Strong hydrogen bonds are rare in biological structures since they are very rigid and not easily broken (D…A distance less than 2.5 Å and a D-H…A angle close to 180°, where D and A are the donor and acceptor atoms, respectively) and can hinder processes such as protein folding or unfolding. On the other hand, the salt bridge intermolecular hydrogen bond, N^{+}-H…$^{-}$O=C, which is present in L-α-glutamine, is one of the two strongest intermolecular interactions that exist in biological compounds, the other being the P-OH…O=P hydrogen bond type, which exists in nucleic acids, due to the strong electrostatic component of the interaction, arising from the charged N and O atoms in the zwitterionic molecules (Steiner, 2002).

From our previous work on the amino acids (Moggach *et al.*, 2005a) we generally find that the N-H…O hydrogen bond distances tend to compress to a value no less than approximately 2.65 Å for the before a structural phase transition takes place which relieves the strain on the bonds. This "minimum" value coincides with the shortest N… O hydrogen bond distance reported in the Cambridge Structural Database (CSD) [2.651 Å, (1S, 2R)-cis-1-ammonioindan-2-ol (R)-2-phenylbutyrate, KAPWAZ, Kinbora *et al.*, 2000]. An example of this behaviour is exhibited by L-serine (Moggach *et al.*, 2005a), which undergoes a phase transition at 5.4 GPa as one of the hydrogen bond distances approaches 2.691(13) Å. However, for the current study of L-α-glutamine, no phase transition was observed up to 4.9 GPa, although the various intermolecular interactions shortened considerably, from values of around 2.7-2.8 Å to values close to 2.65 Å. The lengths of these intermolecular hydrogen bonds converge towards essentially the same distance as that exhibited by the single N-H…O intra-molecular interaction in the structure. This bond shows very little variation with pressure, save for a very slight increase in length.

6.2 Experimental study of L-α-glutamine

6.2.1 Crystal growth

L-α-glutamine was obtained commercially from the Aldrich Chemical Company. Colourless crystals, in the form of needles, were grown from a 3:1 water-ethanol solution by slow evaporation at ambient temperature. One small crystal of dimensions 0.1 mm × 0.1 mm × 0.1 mm was selected for the subsequent experiment.

6.2.2 High pressure X-ray crystallography

The high-pressure experiments were carried out using a Merrill-Bassett diamond anvil cell (Merrill & Bassett, 1974), which has a half-opening angle of 40° and was equipped with 600 μm diamond culets and a tungsten gasket. A 250 μm hole was drilled through the gasket in order to accommodate the L-α-glutamine single crystal. A 4:1 mixture of methanol and ethanol was used as a hydrostatic medium. The fluorescence spectrum of a small ruby chip, which was also loaded into the cell, was used to yield all sample pressures. Pressure measurement was carried out by excitation of the ruby *R1* and *R2* fluorescence line emission with a 632.417 nm line from a He-Ne laser and the resulting ruby fluorescence spectrum was detected with a Jobin-Yvon LabRam 300 Raman spectrometer.

Diffraction data were collected on a Bruker SMART APEX diffractometer with graphite-monochromated Mo-Kα radiation (λ = 0.71073 Å). Data collection and processing procedures for the high-pressure experiments were as described by Dawson *et al.* (2004) (Bruker-AXS, 1997-2001; 1999). Data collections were taken in steps from 0 GPa up to a final pressure of 4.9 GPa. Integrations were carried out using the program SAINT (Bruker-AXS, 2003), which resulting in a completeness range of 51.7 to 58.5%, and the absorption corrections were undertaken with the programs SORTAV (Blessing, 1987, 1989) and ABSORB (Angel, 2004). The unit cell dimensions determined were a = 16.023(3), b = 7.7678(18) and c = 5.1004(13) Å at 0 GPa and a = 15.191(8), b = 7.455(5) and c = 4.882(4) Å at 4.9 GPa and the space

group was found to be orthorhombic, $P2_12_12_1$. No discontinuities were observed in unit cell parameters or their first derivatives with pressure, other than what could be expected from a continuous smooth compression and consequently there was no evidence for a structural phase transition.

Refinements were carried out against $|F|^2$ using all data (CRYSTALS, Betteridge *et al.*, 2003). Due to the limitations of the high-pressure data sets, all 1,2 distances were restrained to the values observed in the ambient pressure structure, and all carbon, nitrogen and oxygen atoms were refined with isotropic displacement parameters.

The L-α-glutamine coordinates of Koetzel *et al.* (1973) were refined against these data to yield a conventional *R* factor of 0.1045 for 202 data with $I > 2\sigma(I)$ The R-factor values obtained for the refinements at the different pressures, from 0 to 4.9 GPa, fell in the range of 0.0773 to 0.1483. The aim of the zero-pressure experiment was simply to study the effect of the diamond anvil cell on the quality of the diffracted intensities and the subsequent structural refinement. The resulting crystallographic data can be found in Table 6.1. It was found that the refined structure obtained with the DAC at 0 GPa is consistent [within standard error] with those of Cochran *et al.* (1952) and Koetzel *et al.* (1973) reported at ambient conditions (i.e. mounted on a fibre). However, the position of the hydrogen atoms differ as they were placed geometrically during the refinement procedure and their positions were not refined due to the limited quality of data, with respect to counting statistics and overall completeness. Nevertheless, the H atoms were allowed to ride with the heavy atoms, contributing this way to the refinement of the model. Consequently, the difficulty of locating the hydrogen atoms during the X-ray diffraction analyses led to the identification of hydrogen bonds from the D…A distances (D = donor, A = acceptor) alone and only the distances between the nitrogen and oxygen atoms involved in the interactions will be discussed. The values for these distances at 0 GPa agree with those reported by Koetzel *et al.* (1973) at room temperature.

Pressure	**0 GPa**	**0.1 GPa**	**0.8 GPa**	**1.4 GPa**	**2.7 GPa**	**3.6 GPa**	**4.9 GPa**
				Crystal data			
Chemical formula	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$	$C_5H_{10}N_2O_3$
M_r	146.15	146.15	146.15	146.15	146.15	146.15	146.15
Cell setting,	Orthorhombic	Orthorhombic	Orthorhombic	Orthorhombic	Orthorhombic	Orthorhombic	Orthorhombic
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, b, c (Å)	16.023 (3), 7.7678 (18), 5.1004 (13)	15.992 (2), 7.7558 (12), 5.0941 (9)	15.879 (6), 7.705 (3), 5.084 (2)	15.679 (11), 7.628 (6), 5.023 (5)	15.450 (8), 7.550 (6), 4.972 (5)	15.328 (7), 7.497 (5), 4.941 (4)	15.191 (8), 7.455 (5), 4.882 (4)
V (Å^3)	634.8 (2)	631.83 (17)	622.0 (4)	600.8 (9)	580.0 (7)	567.8 (6)	552.8 (6)
Z	4	4	4	4	4	4	4
D_x (Mg m^{-3})	1.529	1.536	1.561	1.616	1.674	1.710	1.756
Radiation type	Mo $K\square$	Mo $K\square$	Mo $K\square$	Mo $K\square$	Mo $K\square$	Mo $K\square$	Mo $K\square$
No. of reflections	348	551	347	313	325	384	190
σ range (°)	2.5–23.8	2.5–23.2	2.6–22.5	2.6–23.6	2.6–21.3	2.7–23.1	5.4–22.5
μ (mm^{-1})	0.13	0.13	0.13	0.13	0.14	0.14	0.15
Temperature (K)	293	293	293	293	293	293	293
Crystal form, colour	Needles, colourless						
				Data collection			
Diffractometer	Bruker AXS						
Data collection method	ϕ & ω scans						
Absorption correction	Multi-scan (based on symmetry-related measurements)						

Pressure (*Cont.*)	**0 GPa**	**0.1 GPa**	**0.8 GPa**	**1.4 GPa**	**2.7 GPa**	**3.6 GPa**	**4.9 GPa**
T_{min}, T_{max}	0.407, 1.000	0.507, 1.000	0.657, 1.000	0.487, 1.000	0.557, 1.000	0.308, 1.000	0.722, 1.000
No. of measured, independent and observed parameters	337, 326, 202	336, 324, 205	341, 330, 206	305, 295, 202	281, 271, 190	281, 270, 196	292, 283, 182
Criterion for observed reflections				$I > 2\sigma(I)$			
R_{int}	0.184	0.224	0.185	0.155	0.140	0.144	0.214
σ_{max} (°)	23.2	23.3	23.4	23.3	23.3	23.4	23.2
Range of h, k, l	$-17 \rightarrow h \rightarrow 17$	$-17 \rightarrow h \rightarrow 17$	$-17 \rightarrow h \rightarrow 17$	$-17 \rightarrow h \rightarrow 17$	$-17 \rightarrow h \rightarrow 17$	$-17 \rightarrow h \rightarrow 16$	$-16 \rightarrow h \rightarrow 16$
	$-8 \rightarrow k \rightarrow 8$	$-8 \rightarrow k \rightarrow 8$	$-8 \rightarrow k \rightarrow 8$	$-8 \rightarrow k \rightarrow 8$	$-8 \rightarrow k \rightarrow 8$	$-8 \rightarrow k \rightarrow 8$	$8 \rightarrow k \rightarrow 8$
	$-3 \rightarrow l \rightarrow 3$	$-2 \rightarrow l \rightarrow 2$	$-3 \rightarrow l \rightarrow 3$	$-2 \rightarrow l \rightarrow 2$	$-3 \rightarrow l \rightarrow 3$	$-2 \rightarrow l \rightarrow 2$	$-3 \rightarrow l \rightarrow 3$
				Refinement			
Refinement on				F^2			
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.105, 0.263, 1.05	0.148, 0.284, 1.16	0.115, 0.249, 1.04	0.107, 0.243, 1.04	0.077, 0.178, 1.02	0.091, 0.188, 1.09	0.114, 0.183, 1.10
No. of relections, no. of parameters	297, 41	267, 41	330, 41	295, 41	271, 41	270, 41	278, 41
H-atom treatment				Mixture of independent and constrained refinement			
Weighting scheme*							
P(1), P(2)	0.106, 4.74	0.00, 11.6	0.0550, 5.43	0.0897, 4.49	0.0622, 1.70	0.344, 3.13	0.00, 3.26
$(\Delta/\sigma)_{max}$				<0.0001			
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.57, –0.54	1.01, –0.76	0.71, –0.76	0.76, –0.55	0.40, –0.39	0.53, –0.54	1.00, –1.05

Computer programs: *SMART* (Siemens, 1993); *SAINT* (Siemens, 1995); *CRYSTALS* (Betteridge *et al.*, 2003); *CAMERON* (Watkin *et al.*, 1996).
*where $w = 1/[\sigma^2(F^2) + (P(1)p)^2 + P(2)p$.

Table 6.1: Crystallographic data for L-α-glutamine at increasing pressures.

6.2.3 Results and discussion

6.2.3.1 Analysis of the crystal structure of L-α-glutamine at ambient pressure

L-α-glutamine crystallises with one molecule in the asymmetric unit and a total of four molecules in the unit cell. The structure in the solid state is characterised by the formation of molecular layers with the L-α-glutamine molecules adopting a herringbone-like arrangement within each layer (Figure 6.1).

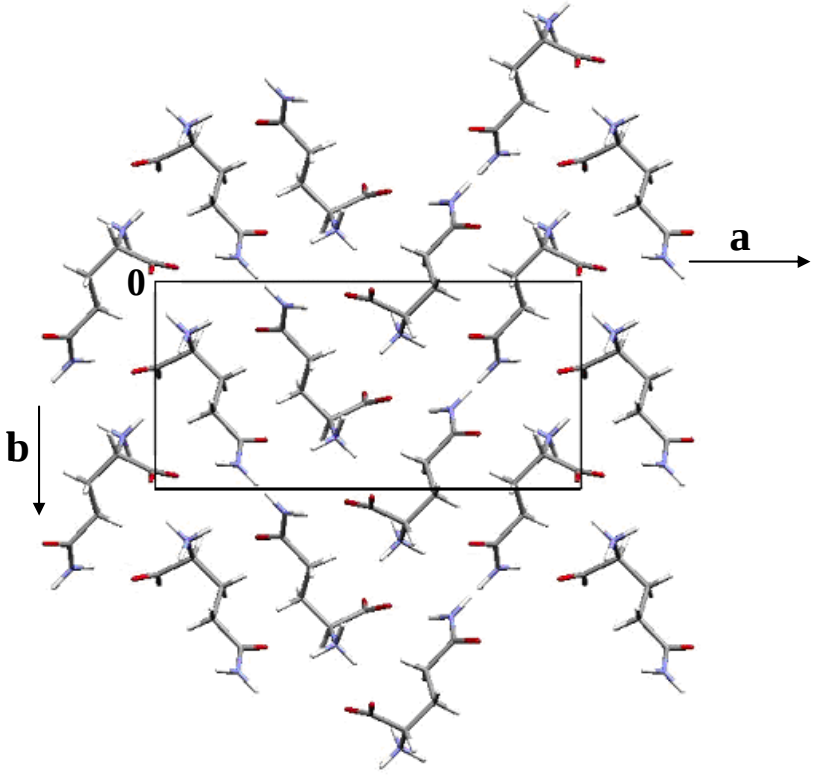


Figure 6.1: View of the packing along the c-crystallographic direction of L-α-glutamine at ambient pressure (0 GPa).

L-α-glutamine exists as a zwitterion state in the solid and aqueous states; this means that the principal hydrogen bond donor group, NH_3^+, cannot, for steric reasons, accept hydrogen bonds, and the principal acceptors, the carboxylate or carbonyl oxygens, have no protons for hydrogen bond donation. There are five different

intermolecular hydrogen bonds present in the L-α-glutamine crystal structure, three formed by the NH_3^+ group and two formed by the NH_2 group Figure 6.2). The distance values for the five intermolecular interactions can be found in Table 6.2, together with the reference values reported by Koetzel *et al.* (1973). Although as already mentioned, there are no strong intramolecular interactions, the distance between the N1 and the O1 atoms is 2.67 Å and the angle formed by N1-H…O1, taking into account that the uncertainty in the position of the hydrogen atom, is only around 90° and consequently by the assignment of Steiner (2002) this can still be considered as a weak hydrogen interaction. Consequently, the NH_3^+ group forms a three-centre or bifurcated hydrogen bond with two different oxygen atoms. The formation of this type of hydrogen bond network was explained in detail by Jeffrey *et al.* (1984).

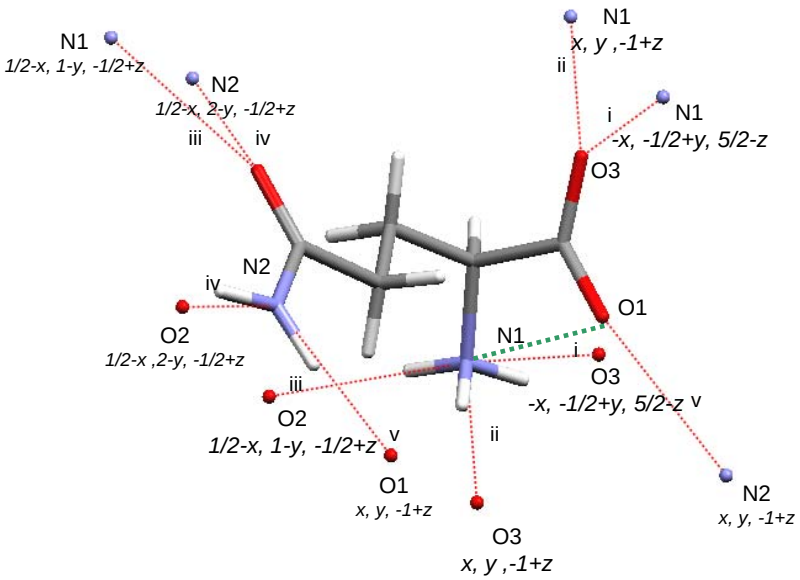
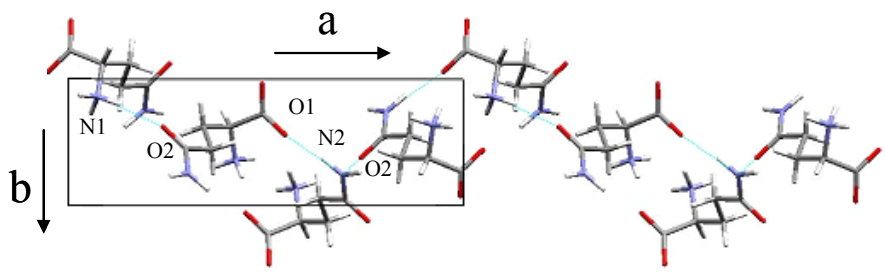


Figure 6.2: View of the hydrogen bonding scheme of one molecule in the crystal structure of L-α-glutamine (the interaction represented in green is the weak intramolecular interaction).

The structure of L-α-glutamine can be also described as corrugated layers of molecules along the *a* axis [Figure 6.3(a)] which are stacked parallel to the 'c' direction [Figure 6.3(b)]. Within each layer, molecules are linked via intra-layer hydrogen bonds: *via* the two interactions formed by the NH_2 groups (N2…O1 and N2…O2), and the N1…O2, formed by the NH_3^+ group [Figure 6.3(a)]. In addition to the intra-layer interactions, there are two inter-layer hydrogen bonds [N1…O3^{i} and N1…O3^{ii}] present in the crystal structure [Figure 6.3(b)]. These two N1…O3 interactions, formed by the NH_3^+ group, stack molecules along the c axis (Figure 6.4). However, due to the proximity of the molecules between layers, the formation of the two inter-layer hydrogen bonds give rise to extra inter-layer interactions, such as N2…O1, which are also present within the layers. The 'c' cell dimension exhibits a value of around 5.1 Å, which is mainly associated with the formation of two head to head hydrogen bonds [N1…O3^{i} and N1…O3^{ii}], which actively participate in the stacking of layers along this direction. This feature is also present in other amino acids, where molecules are arranged in different chain motifs formed by N-H…OOC interactions. In serine (Benedetti *et al.*, 1973) and asparagine monohydrate (Verbist *et al.*, 1972), for example, the axes *c* and *a* are of 5.615(2) and 5.593(5) Å respectively, which are related with the formation of head-to-tail chains along the 'c' and 'a' directions.

In terms of graph-set notation (Bernstein *et al.*, 1995) this rather complex hydrogen bonding scheme can be described in terms of three neighbouring, and coupled, rings. Thus the N1…O3^{i}, N1…O3^{ii} and the N2…O1 form $R_3^2(14)$ rings whereas the N1…O3^{i}, N1…O3^{ii} and N2…O2 interactions form $R_3^3(14)$ rings. Finally, the five different intermolecular interactions can be combined to form $R_3^3(12)$ rings (Figure 6.5).

(a)



(b)

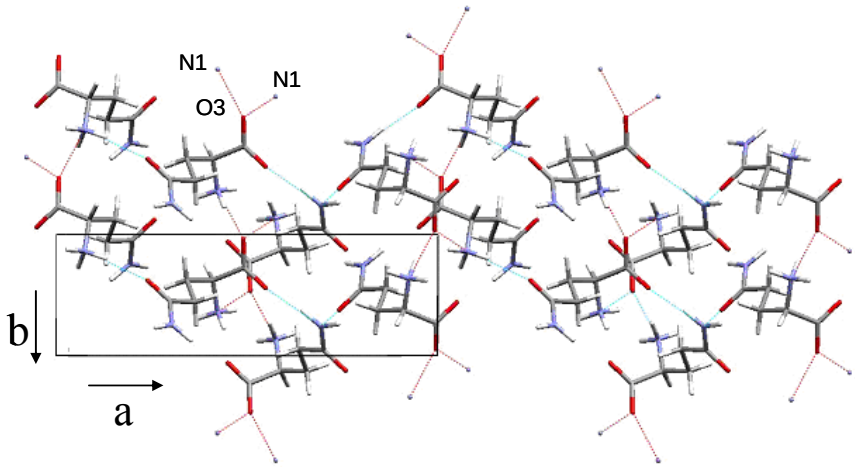


Figure 6.3: View along the *b* axis of (a) a corrugated layer where molecules of L-α-glutamine are linked via the intra-layer (blue) hydrogen bonds to form a corrugated layer parallel to the 'a' direction and (b) layers of L-α-glutamine molecules are stacked along the 'c' direction via the intra-layer (blue) hydrogen.

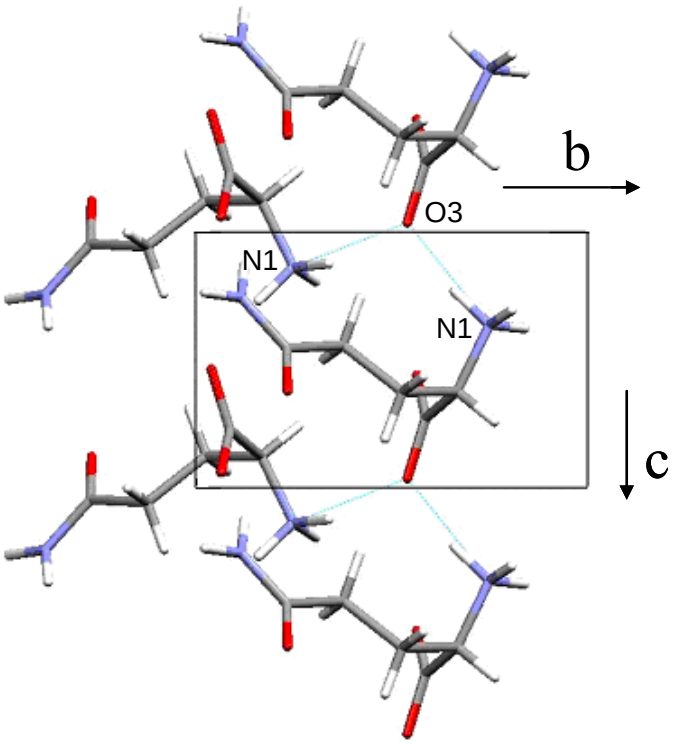
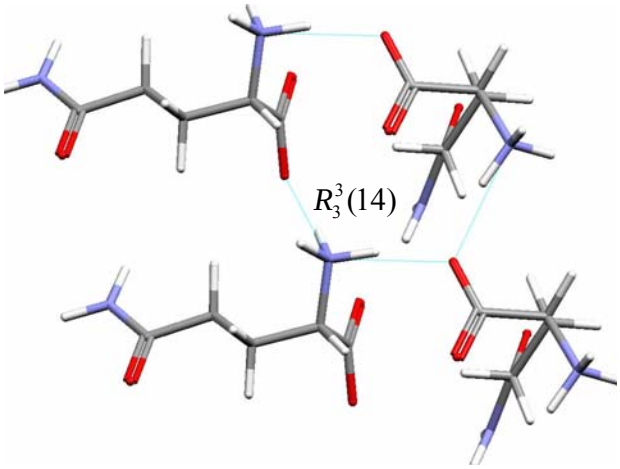
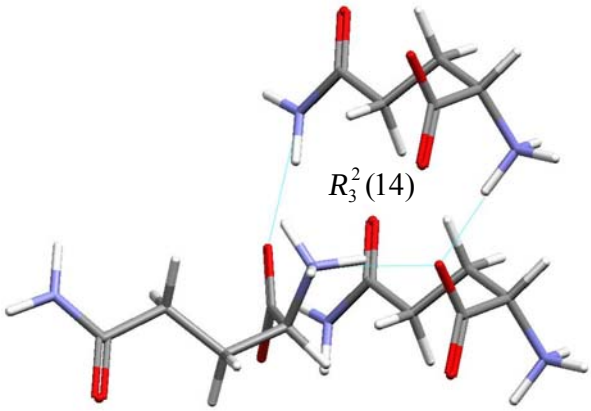


Figure 6.4: View along the *a* axis of the inter-layer hydrogen bonds, which stack the corrugated layers of L-α-glutamine molecules along the 'c' direction.

(a)



(b)



(c)

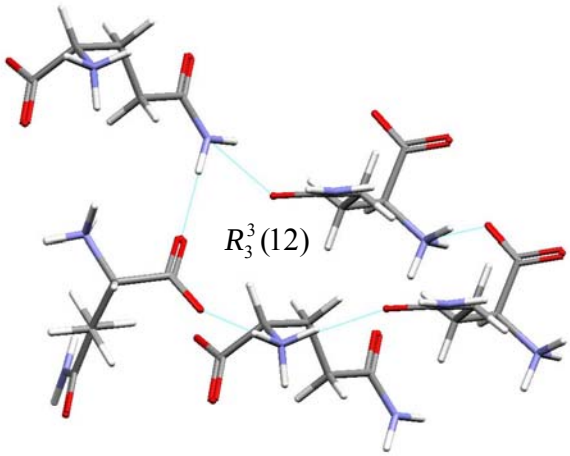


Figure 6.5: View of the three different ring motifs formed by the combination of (a) two [N1…O3^{i} and N1…O3^{ii}], (b) three [N1…O3^{i}, N1…O3^{ii} and the N2…O1] and (c) all five intermolecular interactions in the crystal structure of L-α-glutamine at ambient conditions.

Apart from the strong electrostatic interactions already mentioned, the presence of weak C-H…O interactions in biological systems is also fundamental for their structure. The most prevalent C-H group found in proteins involves the C-α of each amino acid residue. Thus, the large number of such Cα--H…O interactions could affect and influence the primary and secondary structures of proteins as well as their functionality. The Cα-H…O is the most common C-H…O=C type of interaction found in β-sheets, where the C….O distance generally falls in the range 2.91-3.50 Å (with a mean distance of around 3.3 Å) (Desiraju *et al.*, 1999). The second class are C-H…O=C contacts in α-helices with some preference for Cβ-H donors, whereas the third class is composed of interactions to buried polar-side chains. Finally, the fourth class consists of contacts with buried water molecules. In order to complete this study, we have investigated the possibility of weak C-H…O interactions. Thus, in the crystal structure of L-α-glutamine at 0 GPa presented in this work, there are four C-H…O interactions; one formed by the Cα (C2) to the carboxyl group, another one formed by the Cβ (C3) and two formed by the Cγ (C4) of each L-α-glutamine molecule.

6.2.3.2 The effect of pressure on the intra-molecular bond distances and bond angles

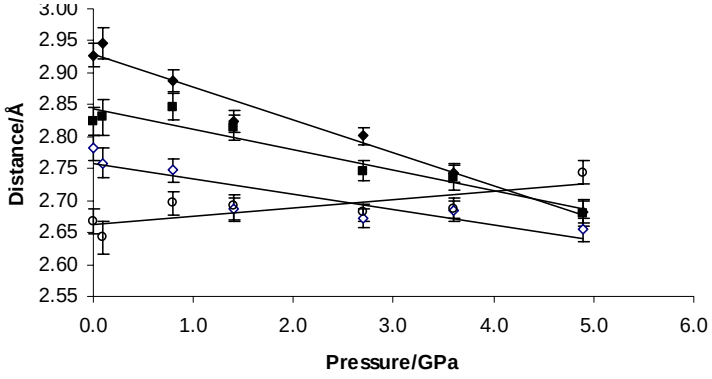
During the crystal structure refinement restraints were applied to the 1,2 distances. Results show that at high pressure, the intra-molecular bond lengths do not change significantly. However, there are some changes in the N1,C2-C1,O3 dihedral angle, which changes progressively from 168.76° to 170.52° respectively, on pressure increase from 0 to 4.9 GPa. This suggests that the molecules tend to increase their planarity with increasing pressure.

6.2.3.3 The effect of pressure on hydrogen bonding

Values of the intermolecular distances with pressure are shown in Table 6.2. On pressure increase to 4.9 GPa, all the intermolecular distances shortened monotonically (Figure 6.6). The two intra-layer N2…O2 and N2…O1 interactions shortened by 4.2% respectively, whereas the N1…O2 interaction changed by 8.3%. As these

hydrogen bonds lie approximately parallel to the a-axis, they will closely correlate to the a-axis compression. The inter-layer hydrogen bonds, N1…O3^{i} and N1…O3^{ii}, shortened by 5.1 and 4.6% respectively from their values at 0 GPa. The reduction of these inter-layer interactions with pressure contribute significantly to the c-axis compression.

(a)



(b)

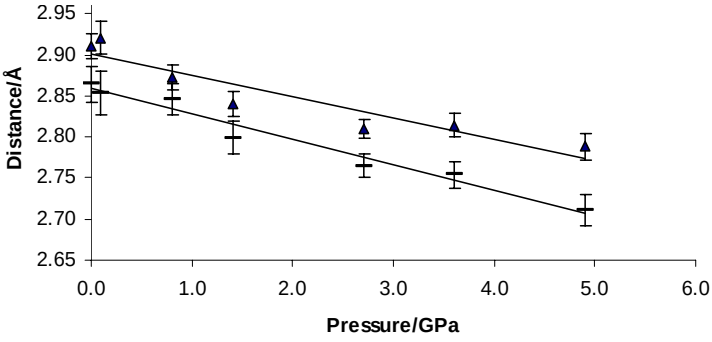


Figure 6.6: (a) Intermolecular and intramolecular hydrogen bond distances (Å) for the three interactions formed by the NH_3^+ group and (b) intermolecular hydrogen bond distances (Å) for the two interactions formed by the NH_2 group, versus pressure (GPa) in the crystal structure of L-α-glutamine. [◊-(N1…O3)^{i}, ■-(N1…O3)^{ii} , ♦-(N1…O2)^{iii}, ▲-(N2…O2)^{iv}, ▬-(N2…O1)^{v}, ○- (N1…O1, intra)]. The lines are fits through the points.

Pressure/ GPa	**Koetzel *et al.* (1973)**	**0**	**0.1**	**0.8**	**1.4**	**2.7**	**3.6**	**4.9**
D(N1…O3)[i]/ Å	2.772(3)	2.782(19)	2.759(24)	2.747(18)	2.686(18)	2.672(14)	2.684(16)	2.654(19)
D(N1…O3)[ii]/ Å	2.866(3)	2.824(22)	2.831(28)	2.847(21)	2.814(20)	2.746(16)	2.736(19)	2.679(20)
D(N1…O2)[iii]/ Å	2.948(3)	2.927(18)	2.946(24)	2.888(17)	2.824(18)	2.801(13)	2.743(15)	2.683(19)
D(N2…O2)[iv]/ Å	2.937(3)	2.911(15)	2.921(20)	2.873(15)	2.840(15)	2.810(12)	2.814(14)	2.788(16)
D(N2…O1)[v]/ Å	2.911(3)	2.864(21)	2.854(27)	2.846(19)	2.799(20)	2.765(14)	2.754(16)	2.744(18)
D(N1…O1)/ Å	2.689	2.668(18)	2.643(25)	2.697(18)	2.691(17)	2.682(13)	2.688(15)	2.744(18)

i) *-x, -1/2+y, 5/2-z*; ii) *x, y ,-1+z*; iii) *1/2-x, 1-y, -1/2+z*; iv) *1/2-x, 2-y, -1/2+z*; v) *-x, 1/2+y, 3/2-z*

Table 6.2: Distances (Å) showing the reduction of the hydrogen bond interactions with pressure for L-α-glutamine inside the diamond anvil cell, from 0 to 4.9 GPa. The values for the intermolecular distance reported by Koetzel *et al.* (1973) at ambient conditions.

Pressure / GPa	**0**	**4.9**
(C…O) interaction	**D(C…O) / Å**	
Cα(C2)…O1	3.54(2)	3.38(2)
Cβ(C3)…O2	3.44(3)	3.07(3)
Cγ(C4)…O3^{i}	3.43(3)	3.22(2)
Cγ (C4)…O3^{ii}	3.518(18)	3.254(18)

Table 6.3: Donor-to-acceptor atom distances (Å) for the four most significant C…O intermolecular interactions present in the crystal structure of L-α-glutamine. These values arose from the data collected on a crystal inside the diamond anvil cell, at 0 GPa.

As can be anticipated there is a clear correlation between the strength of the hydrogen bonds and their compressibility. A clear example of this is the 8.3% reduction in the length of the N1…O2 interaction, which is also the weakest in terms of hydrogen bond geometry. Similar results were found for other amino acids, such as L-α-serine, where one of the N…O distances changed from 2.887(4) to 2.691(13) Å (6.8%) over 4.8 GPa (Moggach *et al.*, 2005a). Finally, the weak intra-molecular, N1…O1, interaction increased its distance by 2.8% on compression from 0 to 4.9 GPa. This was an unexpected result, which shows how the molecule needs to slightly change its conformation in order to accommodate all the changes in the crystal structure produced on pressure increment.

The effect of pressure on the four different C-H…O weak interactions was studied, revealing that the different C…O distances reduced their values by different amounts when pressure was increased up to 4.9 GPa (Table 6.3). Thus, the Cα(2)…O1 interaction reduced its original length by 4.5%, the Cβ(3)…O2 interaction decreased its distance by 10.9% whereas the Cγ(4)…O3 interaction decreased its distance by 6.2%. Finally, the Cγ(4)…O3 interaction shortened by 7.5% its distance at ambient pressure. From these results it can be seen how soft these C-H…O interactions are and they are thus inferred to be much weaker than the N-H…O interactions present in the structure.

6.2.3.4 The effect of pressure on the lattice parameters

The variation of the unit cell parameters and unit cell volume with pressure are shown in Table 6.4. As expected, due to the compression of the intra-layer hydrogen bonds, the a axis has the largest change (5%), whereas both the b and c axes are changed by around 4% at 4.9 GPa, due to changes in the inter-layer interactions (Figure 6.7).

The relative compressibilities of the unit cell edges and the unit cell volume are similar to what we have observed previously for other amino acids. For example in L-α-serine (Moggach *et al.*, 2005a), the largest change in the cell parameters is along the 'b' direction (6.2%), which suffers three times the change along the 'a' and 'c' directions (2.6 and 2.1%). The corresponding change in the unit cell volume was of 11% over 4.8 GPa, which is similar to the 13% reduction in unit cell volume found here for L-α-glutamine over approximately the same pressure range.

Pressure (GPa)	***a* (Å)**	***B* (Å)**	***c* (Å)**	***V* (Å^3)**
Koetzel *et al.* (1973)	16.020(10)	7.762(6)	5.119(4)	636.5
0	16.023(3)	7.7678(18)	5.1004(3)	634.8(2)
0.1	15.992(2)	7.7558(12)	5.0941(9)	631.83(17)
0.8	15.879(6)	7.705(3)	5.084(2)	622.0(4)
1.4	15.679(11)	7.628(6)	5.023(5)	600.8(9)
2.7	15.450(8)	7.55(6)	4.972(5)	580.0(7)
3.6	15.328(7)	7.497(5)	4.941(4)	567.8(6)
4.9	15.191(8)	7.455(5)	4.882(14)	552.8(6)

Table 6.4: Lattice parameters obtained in the L-α-glutamine X-ray diffraction experiments between ambient pressure and 4.9 GPa. The values obtained by Koetzel et al. (1973) are also given for comparison.

Finally, we used the program EOSFIT 5.2 (Angel, 2002) to fit the pressure dependence of the unit cell volume of L-α-glutamine to the Birch-Murnaghan equation of state. Given the relatively low number of observations, *k'* was fixed to a value of 4.0 and *k''* was set at –0.01498 in order that the higher-order terms of the equation were eliminated. The refined values for V_0 and k_0 are 635(2) $Å^3$ and 260(11) GPa respectively. The value of V_0 obtained from this least-squares refinement is in good agreement with the measured ambient pressure value [634.8(2) $Å^3$] and the value of k_0 is very similar to that observed in other amino acid systems [e.g. L-α-asparagine monohydrate, 176(9) GPa]. Finally, a plot of the unit cell volume versus pressure, including the Birch-Murnaghan fit, is shown in Figure 6.8.

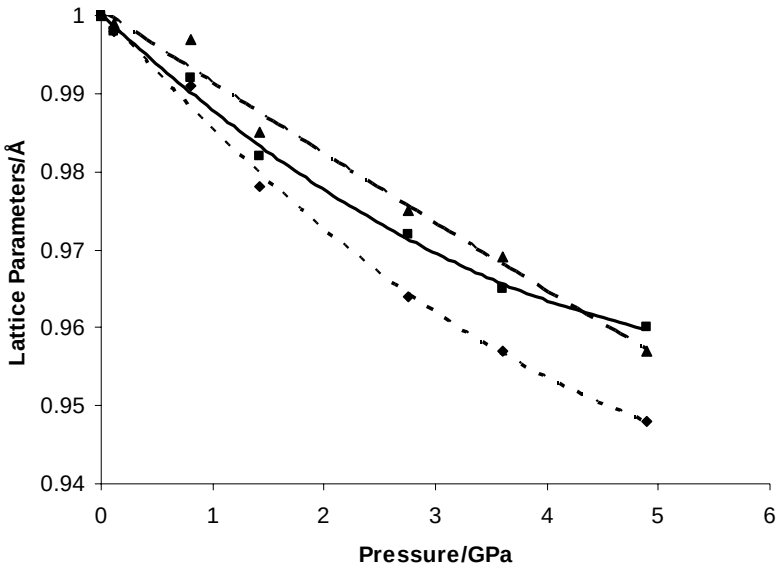


Figure 6.7: Fractional changes in the lattice parameters of L-α-glutamine as a function of pressure (♦, –––= *a* axis; ■, —— = *b* axis; ▲, ••• = *c* axis). The lines are polynomial fits through the points.

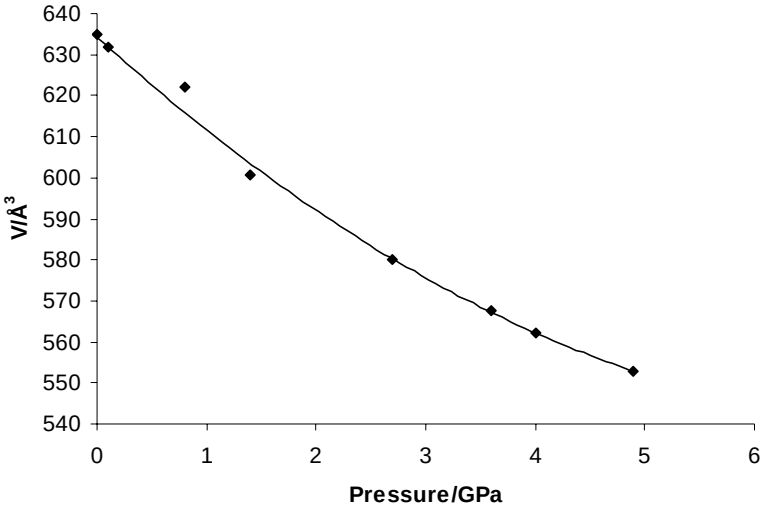


Figure 6.8: Plot of the experimental values and calculated values for the unit cell volume ($Å^3$) of L-α-glutamine. The line is a polynomial fit through the calculated values which have been obtained after a Birch-Murnaghan fit of the experimental data against pressure (GPa) (refining only the values of V_0 and K_0).

6.2.4 Comparison of the ambient pressure and high pressure crystal structures of L-α-glutamine

6.2.4.1 Hirshfeld surfaces

The program CRYSTAL EXPLORER is a recent tool that allows the use of Hirshfeld surfaces to partition crystal space in order to explore packing modes and intermolecular interactions in molecular crystals (McKinnon *et al.*, 2004). We have used this program to visualise the structure of L-α-glutamine at ambient pressure (i.e. 0 GPa) and at high pressure (i.e. 4.9 GPa) in order to make a more detailed comparison between them.

Hirshfeld surfaces (Hirshfeld, 1977) for the structure of L-α-glutamine at ambient pressure and high-pressure are shown in Figure 6.9. It can be seen how the three hydrogen atoms of the NH_3^+ group and the two hydrogen atoms of the NH_2 group actively participate in the formation of hydrogen bonding. This is shown by the

orange-red region on the de surface adjacent to the oxygen atoms, which act as acceptors.

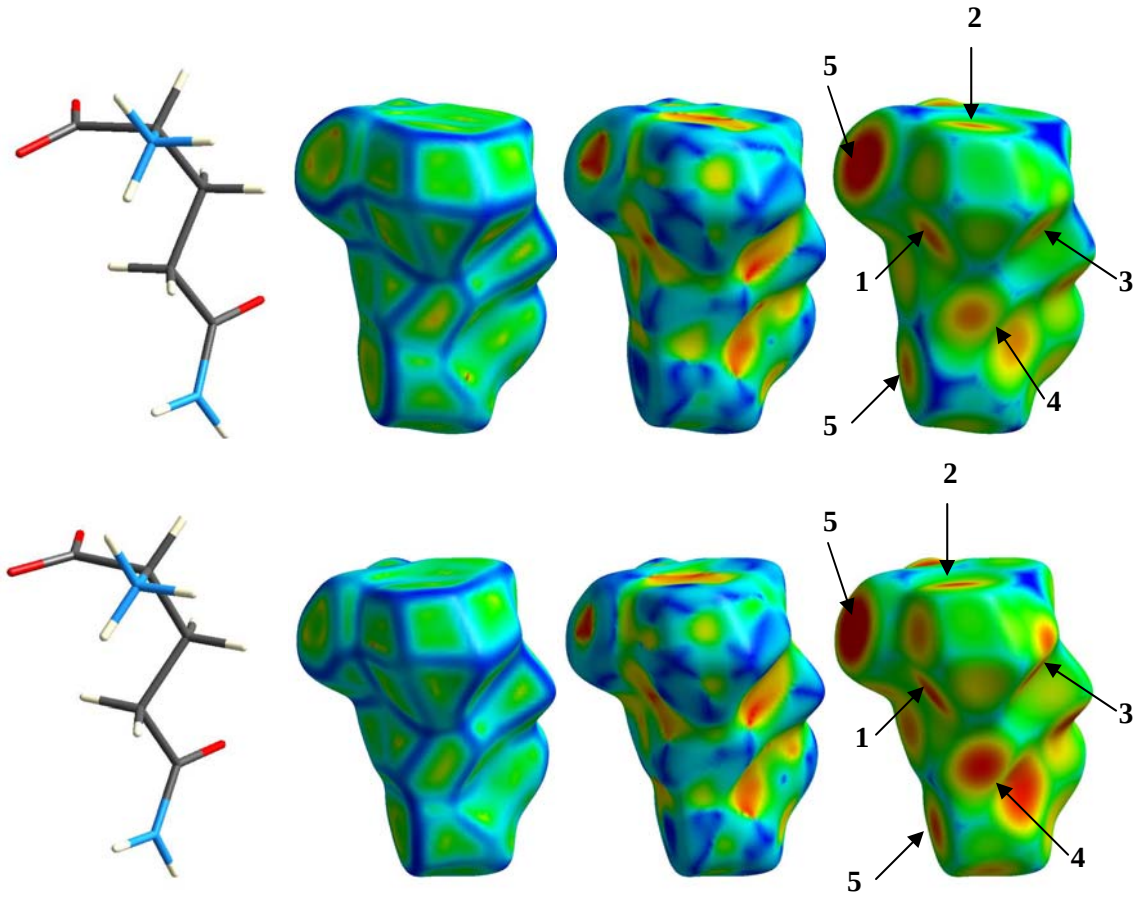


Figure 6.9: Hirshfeld surfaces for the ambient pressure (top) and high-pressure (bottom) structures of L-α-glutamine. Each molecule is shown with the Hirshfeld surface mapped with curvedness (left), shape index (centre) and d_e [right; for this series mapped between 1.0 (red) and 2.0 Å (blue)], where d_e is the distance to the nearest atom centre exterior to the surface. The different interactions are labelled 1-5, as shown in the text: (1) N1…O3^{i}, (2) N1…O3^{ii}, (3) N1…O2, (4) N2…O2, (5) N2…O1.

From Figure 6.9 it is possible to see significant differences in the hydrogen bonding between the ambient and high pressure structures. One of the main differences is the shortening of the hydrogen bonds at high pressure, which is illustrated by an increase in the redness of the contact areas (yellow-orange) in the d_e-surface as well as the formation of other intermolecular interactions (extra red regions in high pressure surfaces), which only become significant when they shorten. In addition to this, the voids 'close up' upon pressure. This can be seen by comparing the

blue areas in the d_e-surface, which are much larger in the ambient pressure structure than those at high pressure. In order to improve the illustration of these differences, the colour settings were changed compared to those used in previous chapters and therefore, the colours in the curvedness surfaces are reversed. These differences are also shown by the fingerprint plots in Figure 6.10.

One of the main differences between the plots is that the voids (upper region of the plots) are more compact at 4.9 GPa than they are at 0 GPa, indicating a more effective packing. The N-H…O intermolecular interactions are shown as spikes in the plot, the area between the spikes corresponds to C-H…O interactions and the spike within the middle area is representative of short H…H contacts. From the plots it is possible to see how the weak C-H…O interaction get stronger when pressure is applied to the crystal structure of L-α-glutamine and how the spikes shortened owing to the hydrogen bond distances decrease.

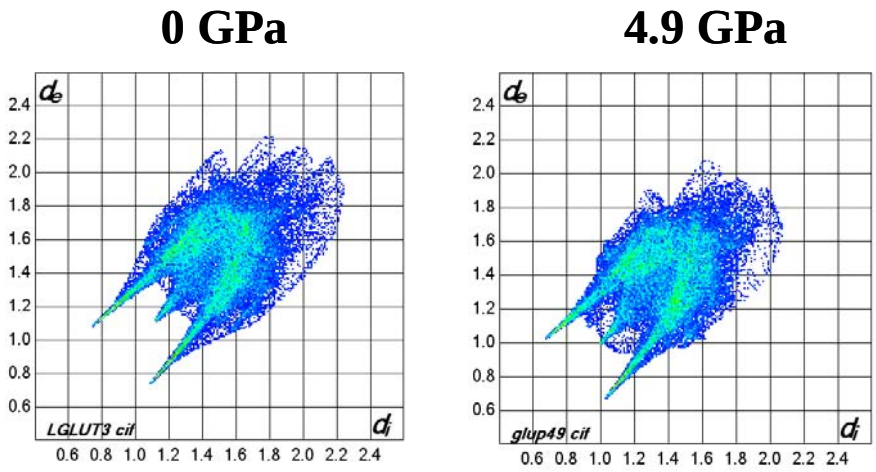


Fig. 6.10: Two-dimensional fingerprint plots for the ambient pressure (left) and high-pressure (right) structures of L-α-glutamine, where d_i, similarly to d_e, is the distance to the nearest atom centre interior to the surface.

6.3 Computational study of L-α-glutamine

6.3.1 Introduction

Following the computational study of α-glycine in Chapter 3, the same computational method, although with more restrictions, was used to investigate the crystal structure of L-α-glutamine. The main reason for carrying out this investigation was to extend the study of a very simple crystalline system, such as α-glycine (simplest amino acid) and L-α-aspartic acid (four carbon-chain molecule), to a much bigger and far more complex structure, such as L-α-glutamine (five carbon-chain molecule), to see if our electrostatic modelling procedure holds for this much more demanding system. This final part of the present chapter aims to not only explore in detail the crystal structure of L-α-glutamine – from its structural features to its energetic properties (i.e., sublimation, lattice and proton transfer energies), but to also find a correlation between the compressibility study described in the first part of this chapter and the individual energies of the hydrogen bonds present in the solid state structure of L-α-glutamine.

The compressibility study of a system gives details about how the crystal structure reacts to pressure variations. In the case of amino acids, many studies have shown how pressure can affect their crystal structure to very different degrees. For example, some structures, such as L-α-glutamine, do not change significantly when pressure is applied to them; others, however, are affected to such a degree that the molecules need to rearrange themselves to form a more stable structure, due to the additional energy arising from the application of pressure. One example of this is glycine (Dawson *et al.*, 2005), which presents various polymorphs of different stabilities between 0 and 6.2 GPa. The experimental compressibility study can give detailed explanations as to how systems modify their structure by tuning the intermolecular interactions with pressure and looking at the resulting changes in bond distances and angles (i.e. geometry). Nevertheless, most changes arise from kinetic and/or thermodynamic alterations, such as the necessity of a crystal structure to lower its packing energy to become more stable than other structures. Unfortunately, no

experiment can give us information about the energetics of the individual intermolecular interactions present in the solid state and therefore computational studies are needed in combination with experimental techniques, in order to provide answers to the many intriguing questions that follow on from the experimental study.

In Chapters 4 and 5 it was shown how the compression of the hydrogen bonds present in the crystal structures of α-glycine and L-α-aspartic acid is in fact directly related to the strength of the interactions. We demonstrated that the weakest interaction, in terms of energy and geometry, will be the interaction with the highest compression, whereas strong interactions are subjected to very small changes. In a simple structure, where the number of hydrogen bonds does not exceed three or four, this relationship is quite straightforward since interactions barely need to compete with others to achieve the conformation which would lead to the most stable crystal structure. However, in the case of more complex structures, such as L-α-glutamine, this situation becomes more complicated, owing to the co-existence of five or more intermolecular hydrogen bonds within the crystal structure. In the case of L-α-glutamine, the unit cell volume is quite large [634.8(2) $Å^3$] compared to α-glycine (310.4 $Å^3$) and L-α-aspartic acid [267.82(16) $Å^3$], which constitutes a problem when trying to reach the basis set convergence, particularly for supercell calculations. In addition to this, the crystal structure of L-α-glutamine is more complex than those for previous amino acids, so that it is not possible to build individual models to accurately reflect the individual H-bonds present in the lattice without the consequent loss of periodicity, which, as shown in Chapters 4 and 5, has a large effect on the resulting strengths of the intermolecular interactions. Therefore, and owing to all the problems already discussed, only the sublimation, lattice and proton transfer energies will be reported for the crystal structure of L-α-glutamine. In terms of intermolecular hydrogen bonding, only the strengths of the individual N-H…O atom-atom pair electrostatic interactions, as gleaned from the V_Coulombic in-house program, will be reported in this chapter. Following the discussion of the results, a comparison between α-glycine, L-α-aspartic acid and L-α-glutamine will be carried out in order to be a step closer towards a profound understanding of the crystal structure of amino acids from an atomistic level.

6.3.2 Results and discussion

The principal aims of this chapter are to once more demonstrate that PW-DFT calculations can achieve reliable estimates for sublimation, lattice and proton transfer enthalpies in the crystal structures of amino acids. Additionally, the electrostatic contribution to the H…O intermolecular interactions will demonstrate, once more, the strengths of the hydrogen bonds in this biological small molecule.

6.3.2.1 Crystal structure of L-α-glutamine at ambient pressure

The starting model to carry out a full quantum mechanical geometry optimisation for L-α-glutamine was taken from the ambient pressure X-ray diffraction study described in the previous part of the present chapter, which is consistent with the neutron diffraction study performed by Koetzel *et al.* (1973). The crystal structure is characterised by the formation of molecular layers with the L-α-glutamine molecules adopting a herringbone-like arrangement with each layer [Figure 6.11(a)]. The structure is stabilised by five N^{+}-H…$^{-}$O=C hydrogen bonds, three of which are formed by the NH_3^{+} group and two by the NH_2 group present in each L-α-glutamine molecule. Note that as the experimental study was part of a high-pressure crystallographic investigation, data were collected in the presence of a diamond anvil pressure cell, which consequently restricted the quantity and quality of diffraction data collected. This had repercussions on the resulting structure refinement, with the hydrogen atoms requiring to be placed geometrically, that is their locations were undetermined by the experimental data.

(a) (b)

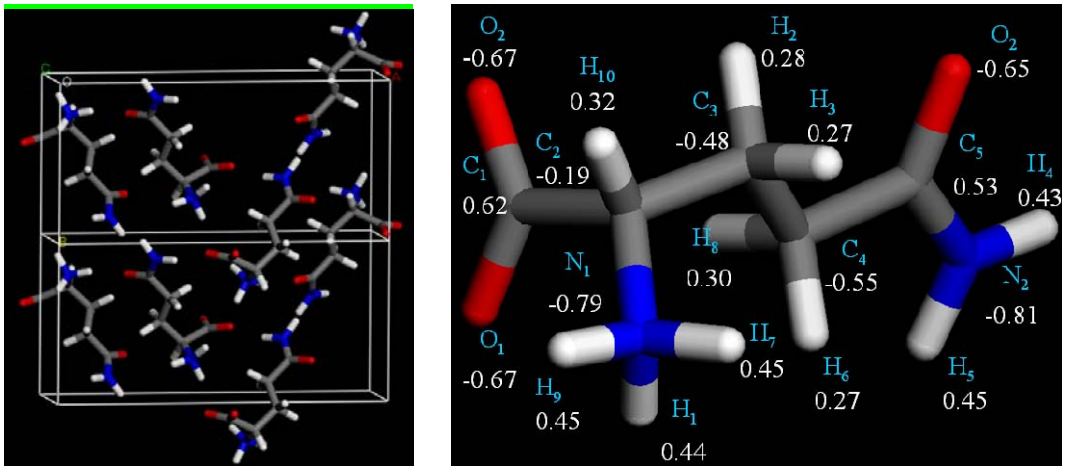


Figure 6.11: (a) The crystal structure of L-α-glutamine. (b) The calculated Mulliken charges for the atoms forming the L-α-glutamine molecule in the solid state structure.

6.3.2.2 Literature energies

L-α-glutamine decomposes under the conditions required for sublimation (Gross *et al.*, 1955); consequently there is very little information regarding the thermodynamic properties of this compound. The only data available in the literature is a reported calculated lattice energy by Volkov *et al.* (2004), who reported values in the range of 42.1 to 56.2 kcal mol^{-1} (176.0 to 234.9 kJ mol^{-1}), depending on the method used to perform the calculation.

6.3.2.3 Geometry optimisation

The results obtained by PW-DFT calculations for the geometry optimisation of the L-α-glutamine structure can be found in Table 6.5, together with the experimental data for comparison. From this it can be seen that in general the computational results are in good agreement with the experimental structure. The differences in cell parameters are just 1.0, 0.8 and 1.0% for the 'a', 'b' and 'c' directions respectively. The cell volume increased by 2.9% with respect to the experimental value. The simulation reported heavy atom positions largely consistent with the input experimental ones, with differences in the bonding parameters of the order of

picometers [e.g. r(C-N)exp = 1.49(1) Å, r(C-N)calc = 1.485 Å]. The locations of the hydrogen atoms, placed geometrically in the experiment, were up-dated quantum mechanically and were found to remain in essentially equivalent positions, with any differences of the order of picometers. In terms of basis set convergence, the geometry optimisation and Mulliken population analysis calculations were carried out using a basis set of 300 eV giving rise to the results previously discussed. Owing to the large unit cell volume of the crystal structure and the supercell models, it was not possible to carry out the geometry optimisation calculations with higher basis sets to look for basis set convergence. The difference in energy per molecule between the unit cell and supercell calculations found for this system was 1.4 kcal mol^{-1} (6 kJ mol^{-1}), which is much larger than that found for α-glycine (0.2 kcal mol^{-1} or 1 kJ mol^{-1}), for example. Nevertheless, it was possible to perform a single point energy calculation on the crystal structure at 500 eV to calculate the Mulliken charges for the L-α-glutamine molecules. Results suggest that the increment of basis set, from 300 to 500 eV, does not cause a significant effect in the calculated Mulliken charges and, therefore, those calculated at 300 eV will be used to carry out the remaining calculations.

Parameters	**Experimental**	**Calculated**		
	Solid	Solid	Supercell (not relaxed)	Supercell (relaxed)
Lattice [Å, °]				
a	16.023(3)	16.1866	17	17
b	7.7678(18)	7.8337	17	17
c	5.1004(13)	5.1509	17	17
α=β=γ	90.0	90.0	90.0	90.0
Z	4	4	1	1
Volume [Å^3]	634.8(2)	653.1	4913.0	4913.0
Space/point group	*$P2_12_12_1$*	*$P2_12_12_1$*	*P1*	*P1*
Geometry [Å, °]				
Dimer 1				
rN1…O3	2.782(19)	2.814	—	—
rN1-H9	1.000	1.059	1.059	1.022
rO3…H9	1.881	1.774	—	—
∠H9-N1-H1	109.6	110.3	110.3	—
∠H9-N1…O3	148.3	166.4	—	—
Dimer 2				
rN1…O3	2.927(18)	2.884	—	—
rN1-H1	0.999	1.050	1.050	—
rO1-H1	—	—	—	1.003
rO3…H1	1.848	1.858	—	—
∠H1-N1-H7	109.5	106.8	—	—
∠H1-N1…O3	164.5	164.5	—	—
Dimer 3				
rN1…O2	2.927(18)	2.973	—	—
rN1-H7	1.001	1.040	1.040	1.023
rO2…H7	2.056	1.948	—	—
∠H7-N1-H9	109.5	109.1	109.1	108.2
∠H7-N1…O2	144.3	168.0	—	—
Dimer 4				
rN2…O1	2.864(21)	2.928	—	—
rN2-H5	0.998	1.028	1.028	1.015
rO1…H5	1.902	1.919	—	—
∠H4-N2-H5	120.1	116.1	116.1	118.9
∠H5-N2…O1	161.4	166.5	—	—
Dimer 5				
rN2…O2	2.911(15)	2.964	—	—
rN2-H4	1.002	1.024	1.024	1.017
rO2…H4	2.094	2.073	—	—
∠H4-N2…O2	137.4	144.1	—	—
Energy [kcal mol^{-1}]	—			
Lattice	56.2*	—	63.7	—
Sublimation	—	—	—	35.1
Proton Transfer	—	28.6	—	—
Total energy [per molecule, eV]	—	-2799.59340	-2796.83305	-2798.07077

Table 6.5: Comparison between the experimental and calculated (PW-DFT) structures for the crystal structure of L-α-glutamine (* from Volkov *et al.*, 2004).

6.3.2.4 Calculated energies

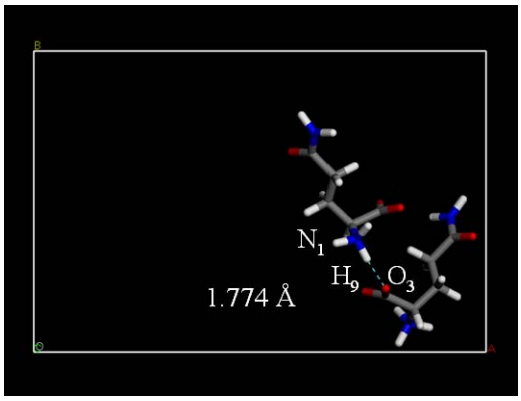
The calculated lattice energy was found to be 63.7 kcal mol^{-1} (266.3 kJ mol^{-1}), which is only 7.5 kcal mol^{-1} (31.3 kJ mol^{-1}) different from the value obtained by Volkov *et al.* (2004) using the CRYSTAL03 simulation package. The value obtained for the sublimation energy of L-α-glutamine was 35.1 kcal mol^{-1} (146.7 kJ mol^{-1}), which is within the range of experimental values found for other amino acids, such as α-glycine [32.6(1) kcal mol^{-1}, 136.3(4) kJ mol^{-1}], alanine [33.0(2) kcal mol^{-1}, 137.9(8) kJ mol^{-1}] and valine [38.9(2) kcal mol^{-1}, 136.2(8) kJ mol^{-1}] (Svec *et al.*, 1965). Finally, the proton transfer energy, calculated from the difference between the sublimation and lattice energies, was found to be 28.6 kcal mol^{-1} (119.5 kJ mol^{-1}), which finds agreement with the data obtained for α-glycine [26.1 kcal mol^{-1} or 109.1 kJ mol^{-1}] during this study, and with other results for amino acids reported in the literature, such as cysteine [26.9 kcal mol^{-1} or 112.4 kJ mol^{-1}] and serine [24.3 kcal mol^{-1} or 101.6 kJ mol^{-1}] (No *et al.*, 1994).

The lattice energy value obtained by PW-DFT can be further tested by performing an Ewald summation, taking the optimised structure and calculated Mulliken charges obtained by the quantum mechanical simulations [Figure 6.11(b)] as input for the program GULP (Gale *et al.*, 2003), as done for α-glycine and L-α-aspartic acid in Chapters 4 and 5. The value thus obtained for L-α-glutamine was 78.1 kcal mol^{-1} (326.5 kJ mol^{-1}), which is considerably higher than the value reported above by PW-DFT calculations. This figure arises purely from a consideration of the electrostatic intramolecular and intermolecular interactions between atoms. The disparity either indicates a deficiency in the calculation of Mulliken charges by the PW-DFT calculations, or points to the fact that a simple electrostatic model alone cannot entirely model the complexity of the L-α-glutamine system. At present, and due to the calculations performed with a higher basis set, we believe the problem to lie with the latter, which means that in all likelihood our computational approach developed for L-α-glycine and L-α-aspartic acid in Chapters 4 and 5 to study (and thus understand) the individual intermolecular interactions present in a molecular crystalline system would be less convincing for this more complicated structure.

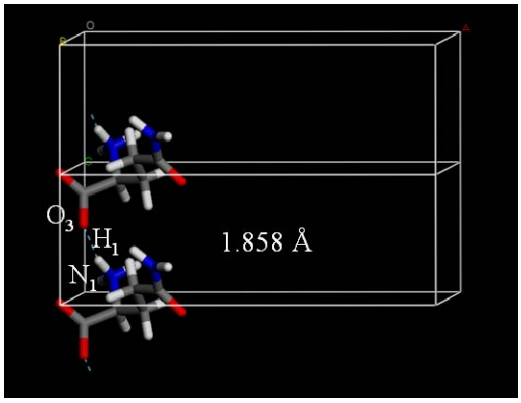
In order to assign energy values to the individual hydrogen bonds present in the crystal structure of L-α-glutamine, five models were constructed as done for previous amino acids, labelled dimer models 1-5 (Figure 6.12), to obtain each hydrogen bond in isolation. Thus, dimer 1 concerns N_1-H_9…O_3 (1.774 Å, 166.4°) and forms molecular chains parallel to the (1 1 0) planes [graph set notation *C(2)*]. Dimer 2 is formed by two molecules of L-α-glutamine linked by the N_1-H_1…O_3 interaction (1.858 Å, 164.5°), which also forms *C(2)* molecular chains along to the (0 1 1) planes. Dimer 3 describes the hydrogen bond N_1-H_7…O_2 (1.948 Å, 168.0°), which forms *C(2)* chains parallel to the (1 0 1) planes. Dimer 4 relates to the N_2-H_5…O_1 interaction (1.919 Å, 166.5°), which forms *C(2)* molecular chains parallel to the (1 0 0) planes. Finally, dimer 5 concerns N_2-H_4…O_2 (2.073 Å, 144.1°) and forms molecular *C(2)* chains along the (0 1 1) planes. Although all the interactions present periodicity in the crystal structure, due to difficulties encountered when constructing the models for the calculations (other interactions were always present apart from the target hydrogen bond in the selected model), models were constructed without periodicity for all the hydrogen bonds.

Akin to the α-glycine structure, we can relate the geometry of the five different hydrogen bonds present in the L-α-glutamine structure with their expected individual energies. On the basis of geometry alone, we would expect the energy of the hydrogen bond interaction to decrease as follows: dimer 1 > dimer 2 > dimer 4 > dimer 3 > dimer 5 (Jeffrey, 1997). The same classical method as used to calculate the strengths of the individual hydrogen bonds for the α-glycine and L-α-aspartic acid systems was applied to L-α-glutamine. The results for these calculations can be found in Table 6.6.

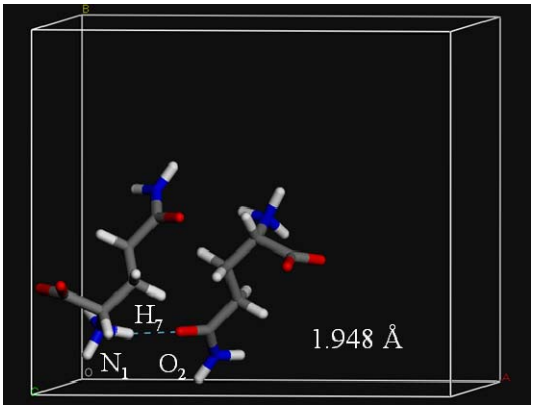
(a)



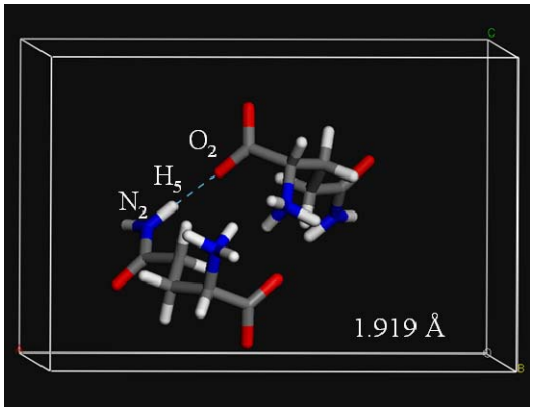
(b)



(c)



(d)



(e)

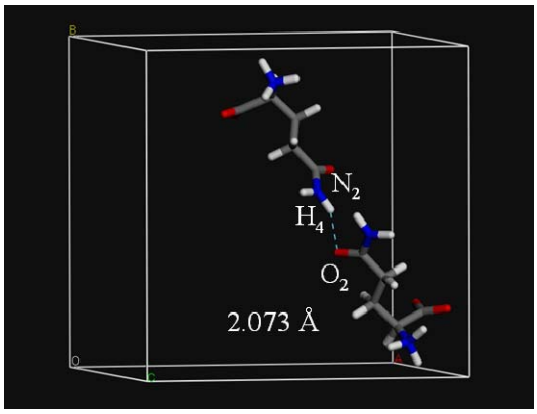


Figure 6.12: The five dimer models used in the calculation of the individual hydrogen bond energies (dimers 1-5, respectively). Note the bond order in models 1, 3-5 is one, whereas for model 2 it is two.

The different energies calculated for the five different individual hydrogen bonds present in the crystal structure of L-α-glutamine can be explained in terms of the electrostatic nature of the interactions. The overall repulsive and attractive energies arising from the atom-atom interactions between molecules in dimer models (as calculated by V_COULOMBIC) are presented in (Table 6.7), along with their sum, which gives total (electrostatic) interaction energies of 32.3, 4.2, 18.3, 22.8 and 4.6 kcal mol^{-1} for aperiodic dimer models 1-5, respectively. It should be remembered of course that these totals include not just the primary H…O interactions, but also the secondary terms arising from the atom-atom intermolecular interactions and, therefore, the overall energies of the D-H…A hydrogen bonds will be influenced by these 'additional' interactions. In addition to this, we also report on electrostatic energies of just the H…O interactions, which were found to be 56.5, 52.7, 49.9, 52.2 and 44.8 kcal mol^{-1} (236.0, 220.4, 208.5, 218.3 and 187.3 kJ mol^{-1}) for dimers 1-5, respectively. Thus, from these results it can be seen how the H…O interaction energies follow the scheme based on the geometry of the hydrogen bonds so that dimer 1 > dimer 2 > dimer 4> dimer 3 > dimer 5.

As can be seen from the results stated above, in this case there is a correlation between the individual hydrogen bond energies and the geometries of the five interactions, which only became apparent for α-glycine upon close examination of the individual atom-atom electrostatic pair contributions. As already mentioned, we believe such an examination for L-α-glutamine will be less reliable. However, for completeness, we report the results obtained here in full.

Model			V_Coulombic
	r(D-H…A) *r* (D…A)	∠(D-H…A) ∠(C-D…A)	Energy/ H- bond
Dimer 1	1.774	166.4	
Aperiodic	2.814	99.7	32.3
Dimer 2	1.858	164.5	
Aperiodic	2.884	116.1	4.2
Dimer 3	1.948	168.0	
Aperiodic	2.973	106.6	18.3
Dimer 4	1.919	166.5	
Aperiodic	2.928	127.0	22.8
Dimer 5	2.073	144.1	
Aperiodic	2.964	143.4	4.6

Table 6.6: Energies obtained per molecule for the L-α-glutamine aperiodic dimers and the resulting energies per hydrogen bond (kcal mol^{-1}) (D – H … A, C-D…A; *r*/Å, ∠/°, D = donor, A = acceptor, C = Carbon atom).

V_Coulombic	**Energy/ kcal mol^{-1}**				
Energy Type	*Dimer 1*	*Dimer 2*	*Dimer 3*	*Dimer 4*	*Dimer 5*
Attractive	*-2585.4*	*-2655.7*	*-2740.9*	*-2510.2*	*-1913.8*
Repulsive	*2553.1*	*2651.5*	*2722.6*	*2487.4*	*1909.2*
Total (hydrogen bond)	*-32.3*	*-4.2*	*-18.3*	*-22.8*	*-4.6*
H…O interaction	*-56.5*	*-52.7*	*-49.9*	*-52.2*	*-44.8*

Table 6.7: Calculated total, attractive and repulsive electrostatic energies (kcal mol^{-1}) for the interactions present in aperiodic dimers 1-5 in the crystal structure of L-α-glutamine. Additionally, the energy of the H…O interaction (kcal mol^{-1}) can also be found.

6.3.3 Compressibility of the L-α-glutamine crystal structure according to computational results

The experimental compressibility study of the L-α-glutamine crystal structure was carried out for the purposes of this thesis and described in detail in the first part of this chapter. When going from 0 to 4.9 GPa, the unit cell parameters decreased their lengths from 16.023(3), 7.7678(18) and 5.1004(13) Å for the 'a', 'b' and 'c' cell dimensions respectively, to 15.191(8), 7.455(5) and 4.882(4) Å, with changes of 5% along the 'a' direction and 4% along the 'b' and 'c' directions. The overall change in the unit cell volume was of 13%, decreasing from 634.8(2) to 552.8(6) Å^3. These changes in the unit cell parameters are a direct consequence of the effect of pressure on the intermolecular interactions present in the crystal structure (i.e. dimers 1-5). Experimental results showed that all the intermolecular interactions decreased their length by 4.8, 5.1, 8.3, 5.4 and 4.2% for dimers 1-5, respectively. The most significant change takes place on the weakest hydrogen bond in terms of geometry (i.e. dimer 3) as it was initially expected, decreasing its length by 8.3%. On the other hand, the smallest compression challenged all expectations, taking place on the weakest interaction in terms on geometry (i.e. dimer 5), which only changed by 4.2%. In general, there is a clear correlation between the effect of pressure (i.e. compressibility) on an H…O interaction and its energy, since the strongest hydrogen bonds exhibited small changes compared to dimer 3. In the case of dimer 5, although it has a very

small energy attributed to it, the interaction participates in the formation of intra-layers of molecules, which play a key role in the stabilisation of the crystal structure, and therefore its compression might require a large amount of energy, which in turn, would destabilise the crystal structure. In terms of the electrostatic energies of the N-H…O interactions, the compressibility-geometry correlation is not that straightforward, most likely due to secondary effects caused by the presence of other atoms in the molecules, apart from those involved in the formation of the hydrogen bonds.

From the combination of the experimental and computational results, a clear correlation between the geometry of the interactions as well as their compressibility, and the energies of the hydrogen bonds can be expected. This was also demonstrated in Chapters 4 and 5 for α-glycine and L-α-aspartic acid. In the case of L-α-glutamine, due to the numerous intermolecular interactions, this correlation weakens, giving in to other stabilisation factors, such as the competition between the interactions for space to form the lowest energy structure. As the system subjected to study becomes more complex, this correlation appears to lose effect and other factors must be taken into account to explain the effect of pressure on the interactions.

6.3.4 Computational

6.3.4.1 Quantum mechanical simulations

6.3.4.1.1 Crystal structure calculations

Calculations on the crystal structure of L-α-glutamine were performed using the CASTEP 2.2 package available through Materials Studio suite of software (Accelerys Inc., 2001-2005). The valence electrons were modelled using a plane-wave basis set expressed at an energy cut-off of 300 eV, which was found to converge the total energies to better than 6 kJ mol^{-1} (~1.5 kcal mol^{-1}) per unit cell in both cases. The electronic core wave function was described using the standard ultrasoft

pseudopotentials available with the software package. The symmetry reduced k-point sets used to sample reciprocal space were generated using Monkhorst-Pack grids (Monkhorst *et al.*, 1977) (dimensions 1 x 3 x 4), giving 4 k-points in the symmetry-reduced first Brillouin zone for L-α-glutamine. The generalised gradient functional PBE was used to model the electronic exchange and correlation.

The initial structures used to perform the geometry optimisations were taken from the ambient pressure X-ray diffraction study of L-α-glutamine, which was presented in the first part of the present chapter. The optimisation of the atomic positions and unit cell parameters were performed on alternate cycles using the BFGS method until the convergence criteria were met (maximum energy change per atom = 2 x 10^{-5} eV, maximum root-mean-square force = 0.05 eV $Å^{-1}$, maximum RMS stress = 0.1 GPa and maximum RMS displacement = 0.002 Å).

A population analysis calculation was performed on the optimised structure to obtain the Mulliken charges for each atom present in the crystal lattice. Since the Mulliken population analysis is formulated in terms of atomic orbitals (Mulliken, 1955), the electronic wave function from a wave is first projected to an atomic orbitals basis set using the scheme of Sanchez-Portal (Sanchez-Portal *et al.*, 1995). The charge spilling parameter for the spin component was found to be 1.17% indicating that a successful charge partitioning was obtained.

6.3.4.1.2 Supercell calculations

These were performed on single, zwitterionic and effectively isolated molecules of L-α-glutamine, for comparison with the energy per molecule in the solid state, and thus allowing the deduction of the total intermolecular interaction energies in the two crystal structures. Zero interaction between the nearest neighbouring cells was obtained by increasing the cell size and observing the change in the total energy. A cell size of 17 x 17 x 17 Å was found to break all intermolecular interactions, without giving rise to overly long computational times. Apart from a reduction in the number of k-points (to one, the gamma point), all other specifications for these calculations

were the same as those applied for the corresponding crystal structure optimisation. The energy of the supercell (no atom relaxation) was compared to the solid state energy of the optimised structure to obtain the lattice energy. In addition to this, an atom-only optimisation calculation was performed on a second supercell to obtain the neutral pseudo-isolated molecule; this energy was then compared to the energy of the optimised crystal structure (per molecule) in order to obtain an estimate of the sublimation energy.

6.3.4.2 Classical mechanical calculations

6.3.4.2.1 Ewald summations

Finally, to obtain the electrostatic contributions to the lattice energy (Ewald sum) the simulation package, GULP (Gale *et al.*, 2003) was used, taking the PW-DFT optimised structure and Mulliken charges as input.

6.3.4.2.2 Coulombic interaction energy calculations

This was obtained *via* an in-house program V_COULOMBIC, written to calculate the atom-atom pair electrostatic energy (Coulombic potential) contributions from a given set of coordinates and point charges.

6.4 Conclusions

The experimental and computational study of the L-α-glutamine structure has been carried out to give a complete ‘picture’ of the amino acid crystal structure. During the experimental investigation, high-pressure X-ray diffraction techniques were used to study the compressibility of the L-α-crystal structure, from 0 to 4.9 GPa. Pressure induces the different intra and inter-layer hydrogen bond distances to shorten by varying amounts and changes ranging between 8.3 and 4.2% were observed. The

hydrogen bond which is reduced the most (8.3%) is N1…O2, followed by changes in N1…O3^{i} (5.1%), N1…O3^{ii} (4.6%), N2…O2 and N2…O1 (4.2%). Consequently, the molecules within layers are pushed together and the layers are compressed along the c axis. This reduction in the hydrogen bond distances can be also described using the Hirshfeld surfaces and fingerprint plots presented in this work. Additionally, significant changes occurred on the weak C-H…O interactions, with C…O distances shortened by diverse amounts, from 4.5 to 10.9% of their values at 0 GPa, therefore giving an insight of their soft nature.

The unit cell volume data of L-α-glutamine, from 0 to 4.9 GPa, were fitted to the Birch-Murnaghan equation of state. Thus, values for the V_0 and k_0 were found to be 635(2) Å^3 and 260(11) respectively. The value of k_0 is similar to the value obtained from a high pressure study of L-α-asparagine monohydrate and falls broadly within the range of other hydrogen-bonded systems.

L-α-glutamine does not undergo a phase transition up to a pressure of 4.9 GPa as was also found to be the case for L-α-aspartic acid up to a similar pressure. However, the much simpler amino acid, L-serine, does undergo a phase transition at about 4.8 GPa, to a previously unobserved polymorph (Moggach *et al.*, 2005a). A reason for this can be found in the shape and size of the molecules. For example, the smallest amino acid, glycine, which forms simple hydrogen bond networks, presents several phase transitions [Boldyreva *et al.* (2003, 2004a, 2004b), Dawson *et al.* (2005)]. However, for larger amino acids such as L-α-glutamine and L-α-aspartic acid, this is not found to be the case. Due to the flexibility of the molecules, the presence of several hydrogen donor and acceptor atoms and the necessity to form multiple hydrogen bonds to hold the crystal structure together, these larger molecules find difficulty in achieving a minimum in their packing energy in order to go through a phase transition, probably due to kinetic impediments. Nevertheless, in the compression of the structure of L-α-glutamine, the “limit” distance was reached but no phase transition happened. This question of how much pressure is required to drive the structure to a new polymorph is thus still open.

Computational results showed that, as it was already proved for α-glycine and L-α-aspartic acid, the individual energies of the hydrogen bonds which form the crystal structure of L-α-glutamine are indeed correlated with the geometry of the interactions, although this correlation is much weaker than that found for simpler amino acids. This might be due to the fact that the crystal structure of L-α-glutamine is more complex: molecules are not only arranged in layers but also in a herringbone-like type of packing, and there are five strong hydrogen bond interactions instead of only three or four strong hydrogen bonds. Due to the electrostatic character of the zwitterionic molecules, the proximity of atoms from different molecules give rise to secondary interatomic interactions which affect the overall energy of the N-H…O bonds, whereas the H…O interaction energy presents very similar values for the different interactions.

In terms of the effect of pressure on the crystal structure of L-α-glutamine, there is a clear relationship between the compressibility and energy of the interactions, so that the weakest interactions are those with the weakest geometry and, therefore, the interactions presenting the largest compression, whereas strong interactions (i.e., high energy and strong geometry) exhibit only small changes. Nevertheless, this correlation is not as clear as it was for previous amino acids, due to the competition between hydrogen bonds to arrange the molecules in order to minimise the lattice energy of the crystal structure.

Despite these differences in the crystal structures of α-glycine, L-α-aspartic acid and L-α-glutamine, all the ambient pressure structures present similar sublimation, lattice and proton transfer energies, which are of the same order of magnitude, typical of amino acid structures, as shown in the literature. In addition to this, the individual energies of the hydrogen bonds present in previous amino acid crystal structures are also of the same order of magnitude, ranging from 1 to around 30 kcal mol^{-1} (4 to around 120 kJ mol^{-1}), which agrees with what was found for L-α-glutamine via classical mechanics calculations. Finally, it was demonstrated, one more time, that the computational method suggested to study the crystal structure of α-glycine in Chapter 4 can be successfully used to study other amino acid structures,

such as L-α-aspartic acid and L-α-glutamine, although careful attention must be paid to the results in order to compare the outcome of the calculations.

6.5 References

Allan, D. R., Clark, S. J. (1999). *J. Phys. Rev. B,* **60**, 6328.

Allan, D. R., Clark, S. J., Brugmans, M. J. P., Ackland, G. J., Vos, W. L. (1998). *Phys. Rev. B,* **58**, R11809.

Allan, D. R., Parsons, S., Teat, S. J. (2001). *J. Synchr. Rad.* **8**,10.

Anderson, P. M.; Ramsay, N. K.; Shu, X. O.; Rydholm, N.; Rogosheske, J.; Nicklow, R.; Weisdorf, D. J.; Skubitz, K. M. (1998). *Bone Marrow Transp.* **22**, 339-344.

Angel, R. J. (2002). *EOSFIT 5.2*. http://www.crystal.vt.edu/crystal/software.html.

Angel, R. J. (2004). *J. Appl. Cryst.* **34**, 486-492.

Benedetti, E; Pedone, C; Sirigu, A. (1973). *Gazz. Chim. Ital.*, **103**, 555.

Bernstein J., Davis, R. E., Shimoni, L.; Chang, N. L. (1995). *Angew. Chem. Int. Ed. Engl.*, **34**, 1555-1573.

Betteridge, P. W., Carruthers, J. R., Cooper, R. I., Prout, K.; Watkin, D. J. (2003). *J. Appl. Cryst.* **37**, 410-416.

Blessing, R. H. (1987). *Cryst. Rev.*, **1**, 3-58.

Blessing, R. H. (1989). *J. Appl. Cryst.*, **22**, 396-397.

Boldyreva, E. V. (2003*). J. Mol. Str.* **647**, 159-179.

Boldyreva, E. V. (2004a). *Cryst. Eng.* **6**, 235-254.

Boldyreva, E. V.; Ivashevskaya, S. N.; Sowa, H; Ashbahs, H.; Weber, H. P. (2004b). *Dok. Phys. Chem.*, **396**, 111-114.

Bruker-AXS (1997-2001). *SMART*, Version 5.049-5.059. Bruker-AXS, Madison, Wisconsin, USA.

Bruker-AXS (1999). *GEMINI*, Version 1.01. Bruker-AXS, Madison, Wisconsin, USA.

Bruker-AXS (2002). *SAINT*, Version 6. Bruker-AXS, Madison, Wisconsin, USA.

Cochran, W. & Penfold, B. R. (1952). *Acta Cryst.* **5**, 644-653.

Dawson, A., Allan, D. R., Clark, S. J., Parsons, S.; Ruf, M. (2004). *J. Appl. Cryst.* **37**, 410-416.

Dawson, A; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L. (2005). *Crystal Growth & Design*, **5**, 1415-1427.

Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, New York, pp. 350-363.

Gale, J. D. & Rohl, A. L. (2003). *Molecular Simulation*, **29**, 291-341.

Gross, D. & Grodsky, G. (1955). *J. Am. Chem. Soc.* **77**, 1678-80.

Hirshfeld, F. L. (1977). *Theor. Chim. Acta.* **44**, 129.

Jeffrey, G. A. & Mitra, J. (1984). *J. Am. Chem. Soc.* **106**, 5546-5553.

Jeffrey, G. A. (1997). *An introduction to hydrogen bonding*, Oxford University Press, New York.

Kinbora, K.; Kobayashi, Y.; Saigo, K. (2000). *J. Chem. Soc. (Perkin Trans.)* **2**, 111.

Koetzel, T. F., Frey, M. N., Lehmann, M. S. & Hamilton, W. C. (1973). *Acta Cryst.* **B29**, 2571-2575.

Lozano-Casal, P.; Allan, D. R.; Parsons, S. (2005). *Acta Cryst.* **B61**, 717-723.

Materials Studio (2001-2005).Accelerys Software Inc.

McKinnon, J. J., Spackman, M. A., Mitchell, A. S. (2004). *Acta Cryst.* **B60**, 27.

Merrill, L. & Bassett, W. A. (1974). *Rev. Sci. Instrum.* **45**, 290-294.

Moggach, S. A; Allan, D. R.; Morrison, C. A.; Parsons, S. & Sawyer, L. (2005a). *Acta Cryst.* **B61**, 58-68.

Moggach, S. A.; Clark, S. J.; Parsons, S. (2005b). *Acta Cryst.* **E61**, o2739-o2742.

Monkhorst, H. J. & Pack, J. D. (1977). *Phys. Rev. B*, **13**, 5188-5192.

Mulliken, R. S. (1955). *J. Chem. Phys.* **23**, 1833.

No, K. T.; Cho, K. H.; Kwon, O. Y.; Jhon, M. S.; Sheraga, H. A. (1994). *J. Phys. Chem.* **98**, 10742-10749.

Sanchez-Portal, D.; Artacho, E.; Soler, J. M. (1995). *Solid State Commun.*, **95**, 685.

Skubitz, K. M; Anderson, P. M. (1996). *J. Lab. Clin. Med.* **127**, 223-228.

Steiner, T. (2002). *Angew. Chem. Int. Ed.* **41**, 48-76.

Svec, J. H. & Clyde, D. D. (1965). *J.Chem. Eng. Data*, **10**, 151.

Verbist, J. J.; Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. (1972). *Acta Cryst.* **B28**, 3006.

Volkov, A. & Coppens, P. (2004). *J. Comp. Chem.* **25**, 921-934.

Watkin, D. J. ; Prout, C. K. ;Carruthers, J. R. ; Betteridge, P. W. (1996). *CRYSTALS*, Issue 10, Chemical Crystallography Laboratory, Oxford, UK.

Chapter 7

Conclusions and Future Work

The use of high-pressure X-ray diffraction techniques to study the crystal structures of small organic and biological molecules, such as cyclopropylamine and various amino acids, has been explained in detail in this thesis. They have been used to obtain a fuller understanding of the main structural features of these systems, while focusing on the profoundly important hydrogen bond. Additionally and to complement the experimental research, a wide range of computational techniques have been used to complete the models obtained from the X-ray diffraction experimental work (e.g. locating H-atoms) and to provide a fuller understanding of the energetics of these systems. Thus, different thermodynamic properties, such as the sublimation, lattice and proton transfer energies together with the individual hydrogen bond energies, could be obtained.

In terms of the experimental work, there is a progression in the complexity of the molecules investigated from cyclopropylamine (which is a small organic molecule) towards larger amino acid sytems such as L-α-aspartic acid and L-α-glutamine. The difference between the two latter amino acids is the existence of two different functional groups (R-group): an amide in the L-α-glutamine case, and a carboxylic acid in the case of L-α-aspartic acid. Both structures share similarities, such as the existence of N^{+}-H…$^{-}$O interactions. However, in the crystal structure of L-α-aspartic acid there is an additional O-H…$^{-}$O interaction.

The study of how the crystal structures of cyclopropylamine, L-α-aspartic acid and L-α-glutamine behave under pressure has led to very different results. In the case of cyclopropylamine, the starting crystal structure went through a phase transition at around 1.6 GPa, whereas the structures of the amino acids did not exhibit phase transitions but compressed monotonically. The main reason for this is that the amino acid systems are much more complex, with a higher number of hydrogen bonds. The ambient structure of cyclopropylamine is stabilised by a short N-H…N hydrogen bond and a longer relatively weak N-H…N interaction, which form layers of trimers connected via the strong N-H…N hydrogen bonds. In the case of L-α-aspartic acid and L-α-glutamine, the ambient structures are formed by layers of molecules, which are linked by moderate/strong N^{+}-H…$^{-}$O and O-H…$^{-}$O hydrogen bonds. Due to the

differences in the hydrogen bonding and other possible crystal effects (i.e., other attractive and repulsive forces), the additional energy applied to the crystal structure of cyclopropylamine by increasing pressure is sufficient to cause a phase transition giving rise to a second polymorph at 1.6 GPa. In the case of the two amino acids, this additional energy is insufficient for a phase transition to take place. However, our group reported a number of investigations carried out on the crystal structures of simpler amino acids, such as glycine (Dawson *et al.*, 2005) and serine (Moggach *et al.*, 2005), where phase transitions have taken place at high pressure (0-5 GPa). In all these cases, molecules are less flexible and form simpler structures than the amino acids studied in this thesis, making the energy required to drive the structures through a phase transition easier to achieve.

In order to understand the energetics of these systems a computational investigation was carried out on each of the following molecular systems: α-glycine (a test based on a high-pressure single-crystal X-ray diffraction study pre-dating this thesis), cyclopropylamine, L-α-aspartic acid and L-α-glutamine. A computational method was proposed to investigate the crystal structures of the amino acids, whereas the computational protocol used for the study of cyclopropylamine was taken from Morrison *et al.* (2003). The α-glycine crystal structure was used as a test to probe the new computational method, which was based on the calculation of the sublimation, lattice and proton transfer enthalpies, together with the individual energies of the hydrogen bonds through two different approaches: classical and quantum mechanics. Both computational approaches gave reasonable results, which were in close agreement. When the computational method was used to investigate the structures of L-α-aspartic acid and L-α-glutamine, this agreement did not appear to be as close as it was for the simplest amino acid, α-glycine.

Computational results in the case of cyclopropylamine show how the sublimation energy of the system can be calculated taking the crystal structure of the molecular system as a starting point to obtain the energy of the molecule in the solid-state environment for comparison with the energy of an isolated molecule (i.e., in the gas phase) or how the individual energies of the hydrogen bonds can be summed to

give an estimate of the total sublimation energy of the system, without having to obtain the overall energy of the crystal structure. This is very useful when the system under investigation is too large for a full quantum mechanical calculation, which would be computationally expensive and very time-consuming. The results obtained for the N-H…N hydrogen bonds present in the low-temperature (Phase I) and high-pressure (Phase II) crystal structures of cyclopropylamine show energies of around 15 kJ mol^{-1} (4 kcal mol^{-1}) when the α-bond cooperativity effect is applied and of almost half that energy when the α-bond cooperativity effect is not present in the interactions. Therefore, the importance of this effect on the two different crystal structures of cyclopropylamine was discussed in detail in the corresponding chapter.

The computational investigations carried out in the three different amino acids (α-glycine, L-α-aspartic acid and L-α-glutamine) showed how as the amino acid structure becomes more complex, the study of the energetics of the system becomes less straightforward than that for simpler systems, most likely due to arising secondary interactions. In terms of sublimation, lattice and proton transfer energies, all the calculated values were compared to reference data taken from the literature when possible (e.g., α-glycine) to ensure that the computational method succeeded. In cases where these energies could not be compared to reference values, the orders of magnitude were taken as reference data. All the energies were of the same order of magnitude as those found for other amino acids in the literature. Energies ranged from around 26 to around 35 kcal mol^{-1} (109-146 kJ mol^{-1}) for the sublimation energies, 50 to 69 kcal mol^{-1} (209-288 kJ mol^{-1}) for lattice energies and 30-35 kcal mol^{-1} (125-146 kJ mol^{-1}) in the case of proton transfer energies. The orders of magnitude of these thermodynamic properties shed light on the strength of these zwitterionic systems in comparison with weaker neutral organic compounds, such as cyclopropylamine.

One of the other principal aims of this thesis was to calculate the individual energies of the hydrogen bonds present in amino acid structures, since relatively little is known about the strength of these interactions. Calculated energies for these hydrogen bonds showed common values, ranging between 1 and 30 kcal mol^{-1} (4-125 kJ mol^{-1}), depending on the interaction. The calculated energies for the hydrogen

bonds were surprisingly high, taking into account the classification of the N^{+}-H…$^{-}$O interactions of moderate strength, in terms of geometry (Jeffrey, 1997). Our results suggest that some of these 'moderate' interactions, in fact, present very high energies, only expected in the case of covalent or quasi-covalent (O-H…$^{-}$O) hydrogen bonds. Nevertheless, one must take into account that these amino acids are in aqueous solution in the processes within the cell and, therefore, the strength of the hydrogen bonds will be diminished.

In addition to this, the study of the σ-bond cooperativity effect, which is only present in the solid state, was also performed on two of the amino acids, in order to get a better understanding of the magnitude of this effect. The results of these calculations demonstrate that periodic hydrogen bonds present higher energies than if they were lacking periodicity, and these interactions can strengthen by as much as 30-50%.

Finally, following the compressibility studies of the three amino acids (including α-glycine taken from Dawson *et al.*, 2005), the possible correlation between the geometry, the energy and the compressibility of the interactions was investigated. For α-glycine, it was found that there was a very close correlation whereas in the case of larger and more complex amino acids this correlation is weaker, possibly due to the secondary interactions present in the crystal structures of these amino acids. Nevertheless, this correlation, although weaker, is still present.

In summary, the work presented in this thesis demonstrates how experimental (high-pressure X-ray diffraction) and computational techniques can be used in combination to obtain a fuller understanding of the crystal structure of molecular systems. The experimental method gives us the arrangement of the molecules within the crystal structure while the computational techniques provide us with information both at the microscopic level, such as bond energies, and at the more macroscopic level, such as lattice and sublimation enthalpies.

In terms of future work, the limitations of the experimental and computational methods are continually being pushed back. One of the main limitations is thought to

be the size of the system and the supercell models needed to investigate different intermolecular interactions, which carry drastic consequences in the performance of the calculations, such as the achievement of the basis-set convergence, as seen in Chapter 6. Nevertheless, the study of larger systems, such as proteins at high pressure, is starting to take place due to advances in methodology (Fourne *et al.*, 2001). Furthermore, X-ray diffraction techniques are reaching the stage where simple proteins can be investigated at synchrotron sources, and these are likely to progress to include even larger systems, such as viruses (Fourne *et al.*, 2003, 2004). In comparison with these experimental advances, computational techniques are much more limited by the size of the system and even small peptides can be challenging, whereas even the high pressure X-ray diffraction structural studies of these systems are starting to be performed on a routine basis. Nevertheless, computational techniques have followed a clear progression from classical mechanics towards quantum mechanical methods, and many studies are being carried out on large biological molecules, such as DNA, RNA, proteins or viruses using many different methods, some of them using Monte-Carlo and Molecular Dynamic simulations. In addition to this, the use of computational methods in pharmaceutical studies has led to a new area of research, in which the main investigation lines are the search for drug targets (e.g. the use of a drug molecule to inhibit a biological function) and the study of polymorphism (Houslay *et al.*, 2003; Frame *et al.*, 2003).

Dawson, A.; Allan, D. R.; Parsons, S.; Ruf, M. (2004). *J. Appl. Cryst.* **37**, 410-416.

Fourne, R.; Kahn, R.; Mezouar, M.; Girard, E.; Hoerentrup, C; Prangé, T.; Ascone, I. (2001). *J. Synchrotron Rad.* **8**, 1149-1156..

Fourne, R.; Girard, E.; Kahn, R.; Ascone, I.; Mezouar, M.; Dhaussy, A. C.; Lyn, T.; Johnson, J. E. (2003). *Acta Cryst.* **D59**, 1767-1772.

Fourne, R.; Girard, E.; Kahn, R.; Ascone, I.; Mezouar, M.; Lyn, T.; Johnson, J. E. (2004). *NATO Science Series, II: Mathematics, Physics and Chemistry*, **140** (High-Pressure Crystallography), 527-542.

Frame, M. J.; Tate, R.; Adams, D. R.; Morgan, K. M.; Houslay, M. D.; Vandenabeele, P.; Pyne, N. J. (2003). *Eur. J. Biochem.* **270**, 962-970.

Houslay, M. D. & Adams, D. R. (2003). *Biochem. J.* **370**, 1-18.

Jeffrey, G. A. (1997). *An introduction to hydrogen bonding*, Oxford University Press, New York.

Moggach, S.; Morrison, C. A.; Allan, D. R.; Parsons, S.; Sawyer, L. (2005). *Acta Cryst.* **B61**, 58-68.

Morrison, C. A.; Siddick, M. M. (2003). *Chem. Eur. J.* **9**, 628.

Appendices

APPENDIX A: Lecture Courses and Meetings Attended

PhD degree: Year 1

Courses:

Oct. 2002	*X-ray Crystallography and Computational Techniques* (University of Edinburgh, Edinburgh)
Dec. 2002	*Unix 1* (University of Edinburgh, Edinburgh)
Apr. 2003	*BCA[1] Intensive Crystallography Course (X-ray diffraction)* (The University of Durham, Durham)
Jun. 2003	*Fortran 90* (University of Edinburgh, Edinburgh)
Jun. 2003	*Thesis workshop* (The University of Edinburgh, Edinburgh)

Conferences attended:

Nov. 2002 *BCA Autumn Meeting* (London)

PhD degree: Year 2

Courses:

Apr. 2004	*Scientific Visualisation* (EPCC[2], University of Edinburgh, Edinburgh)
Jun. 2004	*Workshop: "The nature of hydrogen bonding and density functional theory"* (CECAM[3], Lyon, France). Poster presentation

[1] British Crystallography Association
[2] Edinburgh Parallel Computing Centre

Sep. 2004 *MSSC2004-Modelling in Solid State Chemistry (summer school)*
(The Imperial College of London, London). Poster presentation

Conferences attended:

Apr. 2004 *Young Crystallographers Satellite Meeting* (Manchester)
Apr. 2004 *BCA Spring Meeting* (Manchester). Poster presentation.

PhD degree: Year 3

Conferences attended:

Dec. 2004 *BCA Autumn Meeting* (Birmingham).
Apr. 2005 *BCA Spring Meeting* (Loughborough). Poster presentation.
Aug. 2005 *XX Congress of the IUCr*[4] *Meeting* (Florence, Italy). Poster presentation.

[3] Centre Européen de Calcul Atomique et Moléculaire
[4] International Union of Crystallography

APPENDIX B: Publications

"**Pressure-induced polymorphism in cyclopropylamine**". Lozano-Casal, P.; Allan, D. R.; Parsons, S. (2005). *Acta Cryst.* **B61**, 717-723.

"**High-pressure polymorphism of cyclopentanol ($C_5H_{10}O$): the structure of cyclopentanol phase-V at 1.5 GPa**". Moggach, S. A.; Allan, D. R.; Lozano-Casal, P.; Parsons, S. (2005). *J. Synchrotron Rad.* 12, 590-597.

"Controlling ligand substitution reactions of organometallic complexes: Tunning cancer cell cytotoxicity". Wang, F.; Habtemariam, A.; Van der Geer, E. P.L.; Fernández, R.; Melchart, M.; Deeth, R. J.; Aird, R.; Guichard, S.; Fabbiani, F. P. A.; Lozano-Casal, P.; Oswald, I. D. H.; Jodrell, D. I.; Parsons, S.; Sadler, P. (2005). *PNAS*, **102**, 18269-18274.

"The solid state structure of α-glycine: a study by first principles". Lozano-Casal, P. & Morrison, C. A. (2005). In submission process to the *PCCP*[1] journal.

[1] *Physical Chemistry Chemical Physics* journal